

VOLUME 20 NUMBER 1 MARCH 1967

SOCIOMETRY

**A JOURNAL OF RESEARCH
IN SOCIAL PSYCHOLOGY**

Articles by

STEINER and DODGE

OLMSTED

STRAUS

SIMON

KATZ, BLAU, BROWN and STRODTBECK

NETTLER and HUFFMAN

HENRY

JONES and deCHARMS

Published by

THE AMERICAN SOCIOLOGICAL SOCIETY

SOCIOMETRY

EDITOR

LEONARD S. COTTRELL, JR.

ASSOCIATE EDITORS

R. FREED BALES	HERBERT HYMAN
HERBERT BLUMER	IRVING L. JANIS
JOHN A. CLAUSEN	FREDERICK MOSTELLER
LEON FESTINGER	THEODORE M. NEWCOMB
NELSON N. FOOTE	WILLIAM H. SEWELL

EDITORIAL CONSULTANTS

BERNARD M. BASS	PHILIP J. MCCARTHY
EDGAR F. BORGATTA	RICHARD T. LAPIERRE
ORVILLE G. BRIM, JR.	STANLEY SCHACHTER
LAUNOR F. CARTE	HERBERT A. SIMON
FRED E. FIEDLER	SHIRLEY A. STAR
ROBERT N. FORD	FRED L. STRODTBECK
NATHAN L. GAGE	GUY E. SWANSON
HERBERT GOLDSHAMER	E. PAUL TORRANCE
LEO A. GOODMAN	JOHN W. M. WHITING
ALEX INKELES	ROBIN M. WILLIAMS, JR.
WILLIAM W. LAMBERT	ROBERT F. WINCH

EXECUTIVE OFFICER

MATILDA WHITE RILEY

Published four times a year

March, June, September, December

by

THE AMERICAN SOCIOLOGICAL SOCIETY

at The Waverly Press, Mt. Royal and Guilford Aves., Baltimore 2, Md.

Sociometry was founded in 1937 by J. L. Moreno

Communications for the Editors and all manuscripts should be addressed to Leonard S. Cottrell, Jr., Russell Sage Foundation, 505 Park Avenue, New York 22, New York. Business communications should be addressed to Matilda White Riley, The American Sociological Society, New York University, Washington Square, New York 3, New York.

Copyright 1957 by THE AMERICAN SOCIOLOGICAL SOCIETY.

Entered at the Post Office at Baltimore, Maryland, as second-class matter.

SOCIOMETRY

A JOURNAL OF RESEARCH IN SOCIAL PSYCHOLOGY

VOLUME 20 • NUMBER 1 • MARCH 1957

CONTENTS

A Comparison of Two Techniques Employed in the Study of Interpersonal Perception.....	<i>Ivan D. Steiner and Joan S. Dodge</i>	1
Intergroup Similarities of Role Correlates.....	<i>Donald W. Olmsted</i>	8
Anal and Oral Frustration in Relation to Sinhalese Personality	<i>Murray A. Straus</i>	21
The Compensation of Executives.....	<i>Herbert A. Simon</i>	32
Leadership Stability and Social Change: An Experiment with Small Groups	<i>Elihu Katz, Peter M. Blau, Morton L. Brown and Fred L. Strodtbeck</i>	36
Political Opinion and Personal Security	<i>Gwynne Nettler and James R. Huffman</i>	51
Sibling Structure and Perception of the Disciplinary Roles of Parents	<i>Andrew F. Henry</i>	67
Changes in Social Perception as a Function of the Personal Relevance of Behavior.....	<i>Edward E. Jones and Richard deCharms</i>	75
Annual Author Index—Volume 19.....		iii

SOCIOMETRY

EDITORIAL POLICY

Sociometry is concerned with the entire range of interests and problems represented by research in social psychology. It is the policy of the editors to seek those manuscripts for publication which represent the significant research interests of investigators who are concerned with giving the field of social psychology theoretical structure and reporting research which is clearly focused, well designed, and competently conducted.

While social psychology is presently regarded by most as a field with indeterminate boundaries, it has as its central focus the investigation of the processes and products of social interaction at the interpersonal, intrapersonal, intergroup and intragroup levels and the development of significant generalizations therefrom. In keeping with the more general meaning of the name of the journal emphasis will be placed on measurement of social behavior. However, this emphasis does not exclude the acceptability of good articles which must rely upon qualitative materials and analyses.

The editors and editorial consultants can be expected to subject manuscripts to rigorous criticism and screening according to the best standards of scientific research and at the same time avoid a sterile orthodoxy which would stultify the communication of creative ideas at the growing edge of the science. Thus the journal will strive to be flexible in its response to the publication needs of its contributors.

It is the intention of the editors to avoid any tendency toward professional provincialism and to invite contributions from any sector of the scientific community which promise to further the objectives of the journal.

Annual subscription rate \$9.00; special rate to members of THE AMERICAN SOCIOLOGICAL SOCIETY \$4.50. Single issues \$2.50.

Specifications for Manuscript Preparation for Sociometry may be obtained from the Editorial Office.

Back Issues dated 1937 through 1955 should be ordered from the former publisher, Beacon House, Beacon, New York.

Notice of change of subscriber's address should be sent to THE SOCIETY four weeks in advance, indicating the old as well as the new address.

A Comparison of Two Techniques Employed in the Study of Interpersonal Perception¹

IVAN D. STEINER *University of Illinois*

JOAN S. DODGE, *University of Illinois*

Studies of interpersonal perception frequently involve asking subjects (Ss) to rate their associates on a set of scales descriptive of personality characteristics, attitudes, or interests. Sometimes Ss are asked to describe their associates as accurately as possible on these variables (2, 4, 8, 9); sometimes they are asked to put themselves in the place of their associates and respond as they think these associates would (1, 3, 5, 6, 11, 12). These two kinds of instructions provide Ss with different response sets. In the first case, the S is asked to describe the "manifest stimulus value" of his associates; in the second case, the S is asked to describe these associates *as they see themselves*—or, more accurately, to describe their reported self-percepts. Gage, Leavitt, and Stone (7) have suggested that the discrepancy between these two sets of ratings may be regarded as a measure of the degree to which the rater attributes self-insight and frankness to his associates.

The research reported in this paper attempts to answer the following questions: How different are the descriptions yielded by these two kinds of instructions? Are the discrepancies between the two kinds of ratings influenced by S's sociometric preference for the associates he is rating? Does it matter whether the S has had pleasant or frustrating experiences with the person he is rating? Is one set of instructions more likely than the other to lead the S to rate others as being like himself?

METHODS

The Ss in this investigation were 60 female students from an introductory psychology class who volunteered to participate in a study of how people work together in groups. They reported to the laboratory in groups of three, and were set to work building a design with small wooden blocks. Ss were instructed to build any one of four model designs pictured on cards which were placed before them. Each S was provided with 9 blocks, and it was explained that each design required the use of all 27 blocks. The necessity for cooperative effort was stressed, since the object of the "game" was to produce a design using a minimum number of "moves."

¹ This research was partially supported by a grant from the University of Illinois Research Board.

Ss were not permitted to speak to one another, but communicated about their preference for designs through the use of push buttons which caused lights to flash in front of their partners. The push buttons, lights, and model designs were all numbered serially from one to four, permitting each S to express her preference to the other persons. Ss took turns, each adding one block at a time to the design, or correcting a previously placed block, until a perfect replica of one of the model designs had been produced. A more detailed description of the designs and of the experimental procedures is presented elsewhere (10).

After the group had produced a perfect design, the members responded to a questionnaire. Ss were asked which of their two partners they would most like to have remain in their group if they were doing the task again in competition for a one-hundred-dollar prize. They were also asked whether they would like to have one or both partners replaced in the event of such competition. In addition, each S rated herself on nine five-point graphic rating scales. The scales called for self-ratings on interest in classical music, propensity to blame oneself for failures, religiosity, behavioral consistency, sensitivity to frustration, reliance upon own judgments, satisfaction with a followership role, interest in science, and propensity to yield to others' wishes. After making these ratings, Ss were asked to rate each of their partners on the same scales, indicating the answers each partner should have given, if she knew and admitted the truth about herself. Finally, each S rated her partners again, this time predicting each partner's actual answers on each of the nine scales.

Ten of the 20 groups were subjected to a manipulation which has not yet been mentioned. Before these ten groups began work on a design the experimenter threw a switch which rearranged the circuits connecting push buttons with signal lights. As a consequence of this procedure, half the messages sent by persons in these ten groups were inaccurate (e.g., S pressed the button to indicate a preference for design No. 1, but this caused her partner's No. 3 light to flash). There is clear evidence that this made the task more frustrating and difficult. In response to items on the questionnaire, 17 of the 30 Ss who received this treatment said the task was frustrating, whereas only 3 of the 30 Ss for whom the circuits were not rearranged regarded the task as frustrating. Fifteen Ss in the groups for whom the circuits were changed said they would like to have one or both partners replaced if they were to repeat the task; only five Ss in the other groups indicated a desire to have a partner replaced. In response to a check list of criticisms, members of the former groups made an average of 3.07 criticisms of their partners, whereas members of the latter groups made an average of only .63 criticisms. Furthermore, the groups for whom the circuits were rearranged required an average of 58.3 moves to complete a design; the

other groups required an average of only 34.7 moves. In the subsequent discussion, the ten groups for whom the circuits were secretly rearranged will be called the *frustrated groups*, and the ten groups for whom the circuits were not changed will be called the *unfrustrated groups*.

RESULTS

Each S rated each of her partners twice: once attempting to be as accurate as possible in describing her partner's real attributes (realistic ratings), and once attempting to predict her partner's reported self-percept (self-percept ratings). The over-all discrepancy between these two ratings of a partner was computed by summing the discrepancies across the nine graphic rating scales. Thus if a S's two ratings of a partner's interest in classical music were in the second and fourth spaces of the graphic rating scale, this was counted as a discrepancy of two and was added to the discrepancies on the other eight scales to obtain the over-all discrepancy between the two sets of ratings given by S to that partner. Table 1 reports the means of these over-all discrepancy scores for Ss in the two types of groups. It is to be noted that the discrepancies for Ss in the unfrustrated groups are about the same size when rating a chosen partner as when rating an unchosen partner. But for Ss in the frustrated groups, discrepancies are significantly smaller when rating a chosen partner. (The significance of these differences was evaluated by use of the *t* test for correlated means.)

With nine five-step graphic rating scales it is possible for the two sets of ratings given by any S to a specified partner to differ by as many as 36 points. The obtained discrepancy scores ranged from zero to 15, and the mean of all the discrepancy scores was 7.5. These differences are so large that it seems unreasonable to regard the two techniques as interchangeable.

Because each S also rated herself on each of the nine scales, it was possible to compute assumed similarity scores. Over-all discrepancy scores representing the disparity between the S's ratings of herself and each of her

TABLE 1
Mean Discrepancies between Two Sets of Ratings of Work Partners

Ratings Made by:	N*	Mean Discrepancies in Ratings of:		
		Chosen Partners	Unchosen Partners	P. of Differences between Mean Discrepancies
Unfrustrated subjects	26	7.19	7.12	—
Frustrated subjects	27	6.33	9.22	.01
All subjects	53	6.75	8.19	.01

* N's are less than 30 and 60 due to omissions in the data.

TABLE 2

Mean Assumed Similarity Scores Derived from "Realistic Ratings" and "Self-percept Ratings" of Chosen Work Partners

Ratings Made by:	N*	"Realistic" Assumed Similarity	"Self-percept" Assumed Similarity	P. of Differences between Mean Discrepancies
Unfrustrated subjects	28	8.30	9.11	—
Frustrated subjects	26	10.71	8.82	.01
All subjects	54	9.53	8.96	—

* N's are less than 30 and 60 because of omissions in the data.

TABLE 3

Mean Assumed Similarity Scores Derived from "Realistic Ratings" and "Self-percept Ratings" of Unchosen Work Partners

Ratings Made by:	N*	"Realistic" Assumed Similarity	"Self-percept" Assumed Similarity	P. of Differences between Mean Discrepancies
Unfrustrated subjects	27	11.00	9.75	.10
Frustrated subjects	28	11.64	9.11	.01
All subjects	55	11.33	9.44	.001

* N's are less than 30 and 60 because of omissions in the data.

two sets of ratings of a partner were computed as above by summing the nine separate discrepancies. Tables 2 and 3 report the means of these assumed similarity scores. For all Ss considered collectively there is a significant tendency for Ss to think that the unchosen partner will rate herself as being more like oneself than she really is. However, this tendency is most pronounced among Ss in the frustrated groups. Ss in the frustrated groups also think the chosen partner will rate herself as being more like oneself than she really is, but this tendency is not shared by Ss in the unfrustrated groups.

The product-moment correlation between the assumed similarity scores based on data obtained using these two different techniques is .58 for assumed similarity to the chosen partner and .49 for assumed similarity to the unchosen partner. Stated differently, only about 30 per cent of the variance in one set of assumed similarity scores can be explained as due to variance in the other set of scores. Again the two techniques do not appear to be interchangeable.

DISCUSSION

Before examining the implications of the above data, it is well to explore the possibility that the differences obtained in this study are artifacts of the

method employed. In rating their partners, Ss always made the realistic ratings before making the self-percept ratings, and it is possible that the greater assumed similarity in the case of the second ratings is a consequence of their having been made last. If the Ss' first ratings of a partner were very different from their own self-ratings, any feeling on the part of the Ss that their second ratings should differ from their first might almost necessarily lead the Ss to make the second ratings more like their own self-ratings. However, there are two reasons for rejecting this argument. Ss' first ratings of their partners are not really very different from their own self-ratings. The mean assumed similarity score derived from the first ratings was never larger than 11.64, whereas it could have been as large as 36. (Large scores mean little assumed similarity.) If Ss felt a need to make their second ratings differ from their first, the opportunity to make them less like their own self-ratings was twice as great as the opportunity to make them more like their own self-ratings. Furthermore, a need to make one's second ratings different from the first would presumably operate when rating the chosen as well as the unchosen partner and should have been responsible for equal discrepancies in these two cases. But, as Table 1 shows, the mean discrepancy between first and second ratings is significantly larger when Ss in frustrated groups are rating unchosen partners than when rating chosen partners. An explanation based on the assumption of a need to make second ratings differ from the first ratings does not seem adequate to explain the data.

There is also the possibility of an effect due to regression of ratings toward the mean. Second ratings might tend to regress toward the middle of the rating scales due to lack of reliability in the instrument. And, if Ss' ratings of themselves happened to lie at the middle of the scales, this regression effect would cause the second ratings to reflect greater assumed similarity than the first. This explanation must also be rejected. In the frustrated groups, Ss' ratings of themselves deviated from the middle of the scales by an average of 1.05 steps on the five-step scales; their first ratings of their partners deviated by an average of 1.06 steps. In the unfrustrated groups, the comparable figures are .97 and 1.07. The self-ratings and the first ratings of partners are so similar with respect to deviation from the mid-point that any appreciable regression effect could only be expected to decrease the degree of assumed similarity reflected by the second ratings.

Furthermore, regression effects cannot account for a large portion of the discrepancy between the realistic and self-percept ratings. In the frustrated groups the second ratings deviated from the mean of the scales by an average of only .10 steps less than did the first ratings. This means that over all nine scales, regression effects could be held responsible for a total discrepancy between these two scores of only .9 steps. The mean

obtained discrepancy between these two scores was 7.77. In the unfrustrated groups the second ratings deviated from the mean of the scales by only .06 steps less than did the first ratings. In this case, regression effects could explain a maximum discrepancy of .54 steps between these two sets of ratings, and the mean obtained discrepancy was 7.15 steps.

The data of this study indicate that the two commonly used techniques for obtaining information about an individual's perception of others do not always measure the same thing. When Ss are asked to rate others as they really are, they describe persons whom they do not like as rather different from themselves; when asked to rate others as those persons would rate themselves, assumed similarity is greater. When Ss are rating persons whom they like, these two techniques yield approximately equal assumed similarity scores, unless Ss have been subjected to a frustrating experience.

Throughout the data there are indications that the descriptions yielded by these two techniques are maximally different when individuals have been subjected to a frustrating experience. A clue concerning the reason for this may be found in the results obtained when using unfrustrated groups. Tables 2 and 3 suggest that even among unfrustrated Ss it is the ratings of the unchosen partner which are most systematically affected by the experimenter's choice of research technique. And data cited above indicate that a frustrating experience of the sort employed in this study accentuates tendencies to dislike and reject others. Consequently, the fact that a frustrating experience increases the differences between ratings obtained by the two techniques may be explained as a consequence of the effect of frustration on the sociometric relationships within the group. In this connection, it should be noted that all Ss employed in this study were strangers at the time they began their group activity, and that the experimental conditions may very well have prevented the formation of really strong likes and dislikes, such as may be found outside the laboratory. In real life groups, the two techniques might produce results which are even more strikingly different.

The data of this study suggest a basis for reconciling some of the contradictory findings of research concerned with interpersonal perception. Because the two common techniques for eliciting responses do not necessarily measure the same thing, it is not at all surprising that different researchers, using different techniques, have sometimes produced seemingly contradictory findings. Once it is recognized that an individual's perception of another's real qualities need not coincide with his perception of that person's reported self-percept, such inconsistencies need not be regarded as genuine contradictions. Instead, as Gage, Leavitt, and Stone (7) suggest, such discrepancies may be treated as meaningful data which increase our understanding of interpersonal perception.

SUMMARY

This paper compares two commonly employed techniques for obtaining data on interpersonal perception. One technique yields information about a person's perception of another individual's real qualities; the second technique measures his perception of the other individual's self-percept. These two descriptions of another person are often quite different, especially when the associate being described is not liked, and when the descriptions are given after participation in a frustrating group experience. Recognition of the fact that these two techniques are not interchangeable should lead to a clarification of some of the seemingly inconsistent findings concerning interpersonal perception.

Manuscript received: July 16, 1956

Revised manuscript received: October 3, 1956

Ivan D. Steiner

Department of Psychology

University of Illinois

Urbana, Illinois

REFERENCES

1. Bieri, J., "Changes in Interpersonal Perceptions Following Social Interaction," *Journal of Abnormal and Social Psychology*, 1953, 48, 61-66.
2. Cline, V. B., "Ability to Judge Personality Assessed with a Stress Interview and Sound-film Technique," *Journal of Abnormal and Social Psychology*, 1955, 50, 183-187.
3. Cottrell, L. S., and R. F. Dymond, "The Empathic Responses: A Neglected Field for Research," *Psychiatry*, 1949, 12, 355-359.
4. Estes, S. G., *The Judgment of Personality on the Basis of Brief Records of Behavior*, Cambridge, Mass.: Harvard College Library, 1937.
5. Fiedler, F. E., and K. Senior, "An Exploratory Study of Unconscious Feeling Reactions in Fifteen Patient-Therapist Pairs," *Journal of Abnormal and Social Psychology*, 1952, 47, 446-453.
6. Gage, N. L., "Judging Interests from Expressive Behavior," *Psychological Monograph*, 1952, 66, (18), (Whole No. 350).
7. Gage, N. L., G. S. Leavitt, and G. C. Stone, "The Intermediary Key in the Analysis of Interpersonal Perception," *Psychological Bulletin*, 1956, 53, 258-266.
8. Lundy, R. M., W. Katkovsky, R. L. Cromwell, and D. J. Shoemaker, "Self-Acceptability and Descriptions of Sociometric Choices," *Journal of Abnormal and Social Psychology*, 1955, 51, 260-262.
9. Pepitone, A. P., and R. G. Hayden, "Some Evidence for Conflict Resolution in Impression Formation," *Journal of Abnormal and Social Psychology*, 1955, 51, 302-307.
10. Steiner, I. D., and J. S. Dodge, "Interpersonal Perception and Role Structure as Determinants of Group and Individual Efficiency," *Human Relations* (in press).
11. Suchman, J. R., "Social Sensitivity in the Small Task-Oriented Group," *Journal of Abnormal and Social Psychology*, 1955, 52, 75-83.
12. Trapp, E. P., "Leadership and Popularity as a Function of Behavioral Predictions," *Journal of Abnormal and Social Psychology*, 1955, 51, 452-457.

Inter-Group Similarities of Role Correlates¹

DONALD W. OLMSTED, *University of Wisconsin*

This study develops an aspect of the concept of *social role* by an exploratory analysis of similarities among small organized voluntary groups with respect to three role-related variables—psychological involvement, sociometric status, and duration of membership. Some effort is also made to examine certain relationships between group roles and group processes.

There are a variety of definitions and theoretical contexts in which the widely used term, *social role*, appears in social-psychological literature (3, 5, 7, 9, 10, 12). Some implications of the way the term is conceived for present purposes will be indicated briefly.

To say that role *X* occurs in a group appears to imply that: (a) some group member exhibits certain regularities of behavior which are a function of the operation of the group—that is, which are not simply regularities of behavior expressive of the member's "personality," and (b) such behavioral regularities can be distinguished from those associated with other roles in the group, and from behavior which is not a part of any role (10, p. 331). To say that role *X* occurs in *several* groups appears to imply, in addition, that (a) certain similarities of behavior can be observed among the different persons filling corresponding roles in each group, and (b) these inter-group similarities can be distinguished from those associated with other roles that are congruent in the groups, and from broad similarities of behavior not associated with any role.

It seems reasonable to expect that a member regarded as filling a given role in a group: (a) will have a characteristic degree of psychological involvement (13, p. 622); and (b) will tend to have a characteristic *status* within the group (7, p. 77)—that is, he should be distinguishable in these respects from members filling *other* group roles. If, then, essentially the same role is found in a number of similar groups, the members filling *corresponding* roles in the several groups would be expected to exhibit similarities of involvement and within-group status.

In stable, organized groups, two somewhat overlapping processes can be posited as occurring: (a) the persons constituting the membership are

¹ This paper is based on data secured in connection with Contract N8onr66216 between the Office of Naval Research and the University of Minnesota (Elio D. Monachesi, principal investigator), and developed from a technical report, "Attitudinal Correlates of Role-Selection Processes in Organized Groups," by the author and Philip C. Sagi (now of the University of Alabama). The present work was supported by a grant from the Research Committee of the Graduate School, University of Wisconsin.

changing, with some members "retiring" (completing a normal membership period) or withdrawing (cutting short a normal membership period), and new members being added; and (b) some individual members are shifting from one role to another. This means that such real-life groups are usually composed of both old-timers and recruits, and suggests that the role an individual is filling at a given time may be related to the relative length of his membership. These considerations also indicate the desirability of viewing group roles, at least with regard to formalized or organized groups, as fairly stable *group* characteristics which may be activated by different persons at different times (10, p. 329). Longitudinal study of the shifts in personnel filling essentially the same roles—the study of a group process—is thus indicated.

In the study of natural groups in the field, adequate observation of the multitudinous similarities and differences of behavior which are ordered and summarized by the role concept may not be feasible in many circumstances. Direct evidence, sufficient for the accurate inference of roles, may be difficult to obtain. However, it often may be possible to obtain indicants or measures of the role-related attitude and status variables and to view these as reflections of the several roles filled by the members of various groups.

The foregoing discussion sketches a rationale maintaining that, in organized groups, attitudes of involvement, interpersonal status, and length of membership can be viewed profitably as examples of *role correlates*. This rationale stresses the desirability of specifying group roles which have intergroup applicability. The roles or role categories thus arrived at can be expected to be broad and probably of greater sociological than psychological interest, since nuances of intermember relationships in particular groups would undoubtedly be lost to view. At present, there is no generally accepted system of such role categories available. For the present research, a simple dichotomy of group roles, applicable to any formally organized group and exhaustive of the membership, was selected: officer and constituent.

A member filling either the officer or constituent role at a given time is destined to fall into one of four categories, of importance for group operation, by some future date: (a) he may have continued as, or have become, an officer; (b) he may have continued as a constituent; (c) he may have continued his membership until a "normal completion" of membership occurred; or (d) he may have withdrawn voluntarily from the group, presumably for psychological reasons, as a "premature terminator." The difference between the latter two categories, from the standpoint of the successful continuation of the group, is that the "normal completers" are generally provided for by members' expectations and pose no threat to

group solidarity, whereas the "premature terminators" usually constitute some degree of group "casualty" since the overt act of termination is a denial of the presumed benefits of continued association.

If role-assignment or role-selection *processes* are occurring in a group, then role behavior should be predictive, to some extent, of impending changes in who fills what roles; role correlates should be similarly predictive. In this view, for example, those constituents who in the future will be entering into the officer role should begin to take on the characteristics of the present occupants of the officer role, with respect to role-related variables.

HYPOTHESES

Considering only the simplest, or dichotomous, conditions—two group roles, two points in time, and two reasons for cessation of membership—eight *member categories* can be established. These can be abbreviated for convenience as shown below:

Role at Time 1	Circumstance by Time 2			
	Officer	Constituent	Normal Completion	Withdrawn
Officer (O)	OO	OC	ON	OW
Constituent (C)	CO	CC	CN	CW

The basic hypothesis which this approach formulates and tests is that members who are in *corresponding* member categories (OO, OC, etc.) in a number of similar organized voluntary groups will tend to have similar scores at Time 1 on each of the role-related variables (status, involvement, length of membership). Pooling of members into the eight member categories, regardless of actual group affiliation, is implied as a test procedure, and intergroup appropriateness of measures is assumed.

The "basic hypothesis" refers to over-all differences asserted to exist among the member categories. Several additional "working" hypotheses, suggested by the rationale, are needed to ascertain the characteristics of these differences. In establishing the following hypotheses, an attempt has been made to include fairly systematic consideration of comparisons deemed meaningful and important, rather than simply by "broadside" comparison of all the many possible combinations of pairs and triads and tetrads of categories (8, p. 304; 15).

First, considering only the roles at Time 1, it would be expected that:

Hypothesis 1. Officers (O) will have higher status, greater involvement, and longer duration of membership than constituents (C).

The rationale asserts that if group processes are operative, members should show indications at Time 1, on role-related measures, of roles they

are to assume by Time 2. Specifically, constituents who are to become officers should have begun to acquire "officerlike" characteristics. Hence:

Hypothesis 2. On the mean magnitude of measures taken at Time 1, CO will be between O and the other constituents (CC + CN + CW).

The preceding hypothesis asserts, in effect, that impending role changes are anticipated by the role correlates. Is impending termination of membership also anticipated by the role-related variables? If so, then those who will have left the groups by Time 2 should differ, on the measures taken at Time 1, from those who are to continue membership. Those who are about to terminate membership might be expected to indicate less involvement than those who are to continue, and this could be accompanied by a somewhat lower within-group "friendship" status. Furthermore, these effects should be stronger in the case of those who are to withdraw "prematurely"—with the implication of some dissatisfaction—than among those who will have completed their membership term "normally." However, with respect to length of membership, the continuing members can be expected to have had less time as members than those who are to depart for either reason. The foregoing reasoning produces the following hypotheses, with officers and constituents treated separately, as is implied by Hypothesis 1.

Hypothesis 3a. (OO + OC) will have higher status and greater involvement than ON, and ON will have higher status and greater involvement than OW; but (OO + OC) will have shorter duration of membership than (ON + OW).

Hypothesis 3b. (CO + CC) will have higher status and greater involvement than CN, and CN will have higher status and greater involvement than CW; but (CO + CC) will have shorter duration of membership than (CN + CW).

Finally, officers at Time 1 who are to continue as officers until Time 2, and constituents at Time 1 who are to become officers by Time 2, should have higher status and greater involvement at Time 1 than those who are to continue membership in the constituent role. However, there appears to be no reason why these two categories should differ in length of membership. Thus:

Hypothesis 4a. OO will have higher status and greater involvement than OC, but OO and OC will not differ in duration of membership.

Hypothesis 4b. CO will have higher status and greater involvement than CC, but CO and CC will not differ in duration of membership.

STUDY PROCEDURES

Two samples of college-student voluntary organizations were studied. Sample I consists of 293 members of 23 "Task" organizations, or those

having explicit and objective goals consisting of the creation of products or services for the student body (e.g., governing boards), and Sample II consists of 219 members of 16 "No-Task" organizations, or those whose products and services were primarily for internal consumption (e.g., fraternities). The No-Task groups are thus "sociability" or "congeniality" groups. The "samples" constituted all such groups on a university campus which at the time of study fell within an arbitrary size limitation and were judged to be clearly Task or No-Task in avowed function. Questionnaire data, secured in a testing room under research-staff supervision, were secured for about 93 per cent of the claimed membership of the 39 groups. Group size varied from 6 to 26 members. Of the 39 groups, 17 were all male, 5 were all female, and 17 had both males and females; about 70 per cent of the subjects were male.

As a measure of involvement, a member's score on an eight-item Guttman scale of Personal Involvement (PI) was used. An example of the fixed-response items is: "How important is the group to you?" The coefficient of reproducibility was .90 ($N = 512$). On PI, a score of 1 represents maximum involvement, and a score of 9, minimum involvement. As a measure of within-group status, an index of Sociometric Status (SS) was devised, consisting, for a given member, of the number of sociometric friendship choices received, divided by the highest number received by a member of the group (4). SS could thus range from 0 (lowest status) to 1 (highest status). Time in group (TG) is the number of months a member said he had belonged to the group at the time of the questionnaire session.

About six months after the questionnaires were completed, during which interval the summer recess occurred, a follow-up was conducted to determine the category in which individuals belonged at Time 2. Those who were eligible for membership and were enrolled in school at the time of follow-up, but were no longer members, were placed in a "premature withdrawal" (OW or CW) category, while those who were no longer eligible for membership, or were not enrolled in school (graduated, left for military service, etc.) were placed in a "normal-completion" (ON or CN) category.²

RESULTS

Although it was expected that certain differences between the Task and No-Task groups might well appear, these could not be deduced from the

² Strictly speaking, only those who graduate are "normal completers," but those leaving school for other reasons (e.g., military service, low grades, economic pressure) were classed as normal completions on the assumption that all "other reasons" were not importantly related to the within-group processes under study. Further breakdown would have complicated the procedure and would have led into the thickets of motivational analysis.

TABLE 1

Means of Member Categories on Three Role-Related Measures, by Task and No-Task

Member Category	N	Sociometric Status (SS)	Personal Involvement (PI)*	Time in Group (TG)†
Task Groups				
OO	6	.57	5.2	9.2
OC	3	.92	3.7	13.0
ON	43	.62	4.7	16.8
OW	12	.60	5.9	17.4
CO	29	.52	4.9	13.2
CC	30	.32	4.7	6.0
CN	128	.29	5.7	12.0
CW	42	.25	6.2	12.0
Task	293	.38	5.4	12.4
No-Task Groups				
OO	3	.37	6.7	17.7
OC	25	.61	4.6	22.4
ON	14	.61	5.4	23.3
OW	3	.59	7.0	30.0
CO	22	.54	4.2	13.8
CC	66	.44	5.8	13.7
CN	72	.40	5.7	20.5
CW	14	.19	6.1	14.6
No-Task	219	.45	5.5	18.0

* On PI, higher scores indicate *less* involvement.

† TG scores not available for 18 members.

rationale, so the hypotheses were posed identically for both samples. The number of members in each of the eight member categories, and mean scores on the three measures, are shown for each sample in Table 1. The small number of members falling in certain of the categories unfortunately meant that some tests of the hypotheses would probably have indeterminate results. It can be noted also that the relatively large proportion of members in the normal-completion categories (ON and CN) certifies to the continuing problem of member replacement in such groups (11). Over the six-months interval, 58 per cent of the members of the Task groups, and 39 per cent of the No-Task members, concluded their membership "normally." In the Task groups, 19 per cent of the members withdrew "prematurely" over the interval, compared to 8 per cent of the No-Task members.

The "basic" hypothesis was tested by three analysis of variance tests for each sample. The F-ratios were as follows:

	Task	No-Task
PI	3.10	2.99
SS	9.72	4.04
TG	6.51	3.31

In each case, the indicated probability level is less than .005, interpreted to signify that the eight category means did not come from a common population, or that individuals in a given member category tended to have similar scores on a given measure which differentiated them from individuals in the other categories in a nonrandom fashion.

The numbered hypotheses were tested by analysis of variance and t-tests, with results summarized in Table 2. These hypotheses for the most part specified direction of expected differences, or relative magnitude, so that questions both of directionality and statistical significance of differences are involved.

Hypothesis 1. The direction of difference was as expected in each case—with officers having higher status, greater involvement, and longer membership duration than the constituents, in both samples. However, in No-Task groups, the differences were not statistically significant for SS and PI.

Hypothesis 2. As expected, those constituents who were to become officers (CO) had significantly higher status than other constituents, but less than officers, in both samples. With respect to involvement, an interesting effect appeared in both samples. Although constituents generally gave responses indicating less involvement than did officers (Hypothesis 1), those constituents who were to become officers showed *greater* involvement than officers. This is here taken to represent "superconfirmation" of the hypothesis that the future role would be anticipated in the measure taken at the initial time. In retrospect, it would seem to be a case of the superconformity of those who are to rise to key positions in established social systems; its extent was not fully provided for in the formulation of the hypothesis. On TG, Hypothesis 2 was confirmed for Task groups, while in No-Task groups the CO category had the *shortest* mean duration of membership, contrary to hypothesis (F-ratio significant). That is, in the No-Task groups, the CO members were relatively "rapid risers."

Hypothesis 3a. The direction was as expected in five of six instances, but only one difference was significant. Some of the categories involved had very few members in them (see Table 1). The SS measure did not discriminate significantly among continuing officers, normally terminating officers, and withdrawing officers, in either sample. In both Task and No-Task groups, the officers who were to continue membership, either as officers or constituents (OO + OC), showed greater involvement than

TABLE 2

*Summary, Tests of Selected Differences of Means of Member Categories on Three Measures, by Task and No-Task**

Hypoth Measure	Predicted Relative Magnitude	Group Type	Observed Means	P <
1. SS	O > C	T	.63 > .31	.005
		N-T	.59 > .42	—
PI	O < C	T	5.0 < 5.6	.02
		N-T	5.1 < 5.6	—
TG	O > C	T	16.0 > 11.4	.005
		N-T	22.7 > 16.7	.005
2. SS	O > CO > (CC + CN + CW)	T	.63 > .52 > .28	.005
		N-T	.59 > .54 > .40	.005
PI	O < CO < (CC + CN + CW)	T	5.0 > 4.9† < 5.7	.05
		N-T	5.1 > 4.2† < 5.8	.005
TG	O > CO > (CC + CN + CW)	T	16.0 > 13.2 > 11.1	.005
		N-T	22.7 > 13.8 < 17.1‡	.01‡
3a. SS	(OO + OC) > ON > OW	T	.69 > .62 > .60	—
		N-T	.59 < .61 > .59‡	—
PI	(OO + OC) < ON < OW	T	4.6 < 4.7 < 5.9	—
		N-T	4.8 < 5.4 < 7.0	—
TG	(OO + OC) < (ON + OW)	T	10.4 < 16.9	.02
		N-T	21.9 < 24.3	—
3b. SS	(CO + CC) > CN > CW	T	.42 > .29 > .25	.01
		N-T	.47 > .40 > .19	.01
PI	(CO + CC) < CN < CW	T	4.8 < 5.7 < 6.2	.01
		N-T	5.4 < 5.7 < 6.2	—
TG	(CO + CC) < (CN + CW)	T	9.4 < 12.0	.01
		N-T	13.7 < 19.5	.005
4a. SS	OO > OC	T	.57 < .92‡	—
		N-T	.37 < .61‡	—
PI	OO < OC	T	5.2 > 3.7‡	—
		N-T	6.7 > 4.6‡	—
TG	OO ≈ OC	T	9.2 ≈ 13.0	—
		N-T	17.7 ≈ 22.4	—
4b. SS	CO > CC	T	.52 > .32	.01
		N-T	.54 > .44	—
PI	CO < CC	T	4.9 > 4.7‡	—
		N-T	4.2 < 5.8	.005
TG	CO ≈ CC	T	13.2 > 6.0‡	.005‡
		N-T	13.8 ≈ 13.7	—

* In this table, the relevant "observed means" are presented in the same order as the "predicted relative magnitude" for comparison.

† "Superconfirmation" of hypothesis; see text.

‡ Direction contrary to hypothesis.

those who were to complete normally their membership (ON), and these in turn showed greater involvement than those who were to withdraw (OW), but the null hypothesis of randomness could not be rejected. As expected, in both samples the continuing officers had lower mean duration of membership than those who were to leave the groups, but this difference was significant only for Task groups.

Hypothesis 3b. All directions of difference were as expected, and all but one (PI in No-Task) were statistically significant. That is, continuing

constituents (CO + CC) had the highest sociometric status, the greatest involvement, and shortest membership duration, followed by normal completers (CN) and withdrawing members (CW), in that order.

Hypothesis 4a. It was expected that initial officers who were to continue as officers (OO) would have higher status and greater involvement than officers who were to revert to the constituent role (OC), but the tendencies of the data were *contrary* to both expectations, in both samples (not significantly). It appears that, in these groups, the prospect of "stepping down" from the officer to the constituent role is accompanied by relatively high status and degree of involvement (as measured by SS and PI), compared to prospective continuation as an officer. OO and OC showed no significant differences in either sample in duration of membership; this "confirmation" of expectation is weak, since it may represent an error of statistical inference of "the second kind" (6, p. 64).

Hypothesis 4b. The results for this hypothesis are not clear-cut; the general agreement between Task and No-Task samples on the preceding tests is absent here. The two samples are alike only on SS, with constituents who were to become officers (CO) having higher status than those who were to continue as constituents (CC). In the No-Task groups, the CO category showed greater mean involvement than the CC category, and there was no difference in membership duration—both of which were expected—but in the Task groups, the tendencies were contrary to hypothesis in these respects. With the benefit of hindsight, an "explanation" for the relatively low status and relatively high involvement of those in the CC category in the Task groups can be advanced: their average length of membership at Time 1 was only six months, a mean far lower than that of any of the other categories in either sample (see TG in Table 1). In Task groups, the CC members were still "rookies."

In sum, for the numbered hypotheses, the direction of differences was as expected in 13 of 16 instances for Task groups, and 12 of 16 for No-Task groups. In Task, 12 of 18 hypothetical statements were confirmed to a statistically significant degree, while 8 of 18 such confirmations occurred in No-Task. The two samples were *alike*, in terms of tendencies of differences, in 14 of 18 instances. These summations are of course favorably inflated somewhat, since they are based in part on the use of the same categories in several hypotheses.

DISCUSSION

"Social role" is generally regarded as a unifying and summarizing construct (5, pp. 485ff.; 10, p. 282; 16, p. 317). In terms of research operations, difficulty has been experienced in closely relating concrete behavioral acts to the theoretical abstraction (2). This research seeks one "way out" by

observation of certain regularities in the conditions and products assumed to be associated with role behavior. Thus, individual variations in status, involvement, and duration of membership were shown herein to cluster nonrandomly, for the most part, on the basis of role categorization. Furthermore, these variables were shown to be *predictive* of impending changes in within-group roles, and of reasons for future cessation of membership.

In some research in this area, group roles "exist" by fiat, apparently without awareness on the part of the investigator that within-role similarities and between-role differences in behavior and perception have been assumed. At least two ways of increasing analytic rigor in this regard are evident. One can directly observe behavioral similarities and differences, usually in a laboratory situation, and build up quantitative evidence for the inference of roles, a task which Bales and others have faced (1, pp. 158-164; 14; 2). Or one can take certain formal, functional categories, recognizable by the members, and determine whether the behavior of members in such categories shows such uniformities as would justify calling them roles. This is attempted herein.

An essential feature of the procedure was the pooling of members from a number of similar groups who were in the "same" position vis-a-vis their respective groups. This permitted filling cells of the analytic scheme, which would usually occur only in a fragmentary fashion in a small group taken by itself, and allowed the appearance of over-all tendencies which could not have been observed in a single group. In this connection, it can be noted that even with simple member breakdowns and samples of several hundred subjects each, the size of certain subgroupings was very small, inhibiting adequate tests of hypotheses.

In general, the concern with distinguishing between "premature withdrawal" and "normal completion" as reasons for membership termination has to do with adequacy of group performance or success. Normal completion generally constitutes an organizational fact of life, or objective limiting condition to achievement, about which, presumably, little can be done. Premature or psychological withdrawals represent areas of potential collective modification of behavior, if "improvement" is desired. In some groups, rates of such withdrawal, as indicative of lack of satisfaction or identification, may constitute one criterion of group performance.

The three variables used were chosen for analysis because they represent different *kinds* of variables that might be expected to be related to roles. Involvement (PI) is a measure of the individual's psychological commitment to the group, sociometric status (SS) is one indication of how a member is perceived by others in the group, while time in group (TG) is a non-psychological "background" factor, or an index of exposure to group influence, which in function resembles ecological factors in sociological

theories. The intercorrelations for individuals among the three measures were low in both samples, varying from .06 to .20, yet each measure was related to the member categories, so that the measures would seem to be reflecting largely distinct facets of the role and membership-maintenance behavior. This again points to the unifying attribute of the role concept.

The involvement variable is of particular interest. As Newcomb has pointed out, "role" itself is strictly a *sociological* concept (10, p. 329)—a group's roles are an aspect of that group's structure. On the other hand, the filling of a role by an individual has to do with two intertwined *social-psychological* processes: the prior imagining of the appropriate behavior, or role-taking (16), and the psychological consequences of engaging in such behavior. Both processes may be viewed as aspects of "involvement in a role." If, then, it could be shown that "involvement in a group" occurs in many cases as a product of, or through the mechanism of, involvement in a role, the usefulness of role analysis would be further enhanced.³

In the development of a theory of group interaction, Bales and his associates have made a fundamental distinction between the "task area," with its "adaptive-instrumental" problems, and the "social-emotional area," with its "integrative-expressive" problems (1, p. 10, p. 59; 14). Since the groups studied in the present research were chosen specifically to represent those toward the extremes of a Task to No-Task (or congeniality) continuum of avowed group function, it might have been expected that the No-Task groups would have handled these problem areas differently than would the Task groups, resulting in quite different relations between the involvement and status measures and the member categories. Yet in these natural groups the general *similarity* between Task and No-Task types, with respect to the means of the role and membership-termination categories, seems remarkable. In addition to the results presented in the tests of the hypotheses, it can be noted that, on each measure, the rank order of the eight member categories in the Task sample is quite similar to the No-Task rank order (Spearman *rho* of .90 for SS, .62 for PI, and .74 for TG). The mean length of membership in Task groups was about a year, and in No-Task groups about a year and a half, suggesting that in the groups studied there had been sufficient time and a sufficient degree of interaction for the relevant within-group processes to become stabilized in a similar fashion, despite the differences in avowed function. In addition, the homogeneity of the members of the several groups, and the common cultural setting in which all the groups operated, probably help account for the similarities between the samples.

This homogeneity of membership and of cultural setting—students in recognized organizations on a large university campus—constitute, of

³ This hypothesis is beyond the scope of the present endeavor.

course, definite limiting factors in the generalizability of the results. In particular, the homogeneity of status of these subjects outside their groups is probably great; it is likely that in many groups, operating in different milieu, complex interaction between within-group status and extragroup status would occur.

Any one of the working hypotheses tested herein taken by itself is rather commonplace—the finding, for example, that officers have higher status than other members is hardly startling. Aside from the dictum of science that nothing needs checking more than that which is obvious, the significance of this research lies rather in the indications that: (a) the questionnaire and sociometric responses were meaningfully related to future behavior; (b) variations in three diverse measures, and the behavior they claim to reflect, can be subsumed economically under the role abstraction; and (c) members in the separate groups, pooled into broad categories, exhibited regularities which suggest that similar group processes were operative.

SUMMARY

Similarities in role behavior among organized voluntary groups were studied in terms of three role-related variables: personal involvement, sociometric status, and duration of membership. The analysis was conducted by the assignment of each group member to one of eight member categories, which were established on the basis of three dichotomous criteria. The criteria were: (a) roles—officer or constituent; (b) reason for cessation of membership—normal completion or premature withdrawal; and (c) two points in time, about six months apart.

Two samples of college-student organizations, one consisting of 23 Task groups with 293 members, and the other of 16 No-Task groups with 219 members, were used. Measures, taken at the initial time, were derived from questionnaire responses, and a follow-up determined the appropriate member category at the second time.

It was shown that persons in corresponding member categories in the several groups tended to have similar scores on each of the three role-related variables; that officers tended to have greater involvement, higher status, and longer duration of membership than constituents; that impending changes in role were anticipated by scores on the measures; and that scores on the measures were related to reason for departure from the group. Direction in difference of means of subclasses was as expected in 25 of 32 instances, and 20 of 36 expectations were confirmed by statistical significance. The Task and No-Task samples were similar in tendency in most respects.

The study of role behavior through analysis of role correlates, the de-

sirability of intergroup applicability of role classifications, the relations between roles and group processes, and the unifying characteristic of the role concept appear to be supported by this research.

Manuscript received: July 23, 1956

Revised manuscript received: October 15, 1956

Donald W. Olmsted

Department of Sociology and Anthropology

University of Wisconsin

Madison 6, Wisconsin

REFERENCES

1. Bales, R. F., *Interaction Process Analysis*, Cambridge, Mass.: Addison-Wesley, 1950.
2. Bates, A. P., and J. S. Cloyd, "Toward the Development of Operations for Defining Group Norms and Member Roles," *Sociometry*, 1956, 19, 26-39.
3. Bates, F. L., "Position, Role, and Status: A Reformulation of Concepts," *Social Forces*, 1956, 34, 313-321.
4. Eng, E., and R. L. French, "The Determination of Sociometric Status," *Sociometry*, 1948, 11, 368-371.
5. Hartley, E. L., and R. E. Hartley, *Fundamentals of Social Psychology*, New York: Knopf, 1952.
6. Johnson, P. O., *Statistical Methods in Research*, New York: Prentice-Hall, 1949.
7. Linton, R., *The Cultural Background of Personality*, New York: D. Appleton-Century, 1945.
8. Mosteller, F., and R. R. Bush, "Selected Quantitative Techniques," in G. Lindzey (ed.), *Handbook of Social Psychology*, Cambridge, Mass.: Addison-Wesley, 1954.
9. Neiman, L. J., and J. W. Hughes, "The Problem of the Concept of Role—A Re-survey of the Literature," *Social Forces*, 1951, 30, 141-149.
10. Newcomb, T. M., *Social Psychology*, New York: Dryden, 1950.
11. Sagi, P. C., D. W. Olmsted, and F. Atelsek, "Predicting Maintenance of Membership in Small Groups," *Journal of Abnormal and Social Psychology*, 1955, 51, 308-311.
12. Sarbin, T. R., "Role Theory," in G. Lindzey (ed.), *Handbook of Social Psychology*, Cambridge, Mass.: Addison-Wesley, 1954.
13. Sherif, M., and C. W. Sherif, *An Outline of Social Psychology* (rev. ed.), New York: Harper, 1956.
14. Slater, P. E., "Role Differentiation in Small Groups," *American Sociological Review*, 1955, 20, 300-310.
15. Tukey, J. W., "Comparing Individual Means in the Analysis of Variance," *Biometrics*, 1949, 5, 99-114.
16. Turner, R. H., "Role-taking, Role Standpoint, and Reference-Group Behavior," *American Journal of Sociology*, 1956, 61, 316-328.

Anal and Oral Frustration in Relation to Sinhalese Personality¹

MURRAY A. STRAUS, *Washington State College*

Research in the field of culture and personality has accumulated rapidly in recent years. It has had considerable impact both at the theoretical level in the fields of sociology and psychology, and at the practical level in the fields of psychiatry and pediatrics. The conclusions, often graphically presented, from studies of child rearing in various societies have given new vigor to psychoanalytic theories concerning the implications of gratification of oral and anal functions for personality formation and have had a profound affect on pediatrics.² However, the results of recent attempts to test these conclusions by more rigorous methods have generally been conflicting. No clear-cut pattern of acceptance or rejection has emerged.

Sewell (13) for example, found no significant relation between the infant-training variables considered and the child's personality. Similarly, despite the large number of variables employed and the many correlations computed, Sears *et al.* (12) found relatively few statistically reliable findings and even these few are as often contrary to expectation as confirmatory. Despite the unusual precision of these two research designs, both authors are careful to point out the limitations of the negative conclusions. Sears *et al.* try to fit each of their findings, or lack of findings, into a theoretical matrix useful for future research; and Sewell states:

... it should not be concluded that these results unequivocally refute the claim that infancy is an important period in the development of the individual's personality, or even that the particular training practices studied have a bearing on personality formation and adjustments. To establish the first point would demand both controlled experiments and the study of other aspects of infancy. To establish the second point would demand the corroboration of the results of this study by many and better-designed studies of different culture and age groups (13, p. 159).

¹ This study was made possible by research grants from the University of Ceylon, the University of Wisconsin, and by the 1953 *Britt Foundation Award* for research in social psychology. The data presented are taken from Chapter V of the author's doctoral dissertation (15). Among the many individuals to whom the author is indebted for advice and assistance, special thanks are due to William H. Sewell, the author's major professor; to Bryce F. Ryan, whose support and collaboration in the author's Ceylon research was invaluable; to Emmet E. Baughman, for his stimulating teaching and research in projective methods; and to my wife, Jacqueline H. Straus, whose manifold contributions included the scoring of the Rorschach protocols used in this study.

² The voluminous literature has been well reviewed elsewhere (3, 4, 5, 6, 8, 9, 11, 15) and will not be repeated here.

The present research is in part designed to be one of these needed studies of other culture groups. It is believed to be the first investigation of the relation between infant training and personality in another culture which relates socialization experiences of a *specific* child with estimates of that child's personality as revealed by the Rorschach technique and the California Test of Personality. The use of the correlation method in this study is in contrast to the usual methodology used by anthropologists which consists of determining from informants how children *in general* are brought up in the society under study, and then relating this material by logical analysis to *generalizations* concerning the personality of members of that society. The design of the present study also minimizes biases in sampling, which might result from the informal selection of informants usual in anthropological studies and biases in perception which might result from completely unstructured types of interviewing.

Two key aspects of infant training have been selected for analysis in this paper: feeding practices and elimination training. There are numerous other aspects of infant training which Freudian theory holds to be associated with personality formation and which could have been subject to test in this study. These two were selected because: (a) They are of fundamental importance to development of the two basic psychoanalytic character types, the anal and the oral. (b) It was felt that the greatest controversy has arisen in connection with these two aspects of socialization.

The proponents of the infant-experience theory of personality formation have reacted to the failure of recent research to support their position by claiming that a test of the relationship between a single isolated infant experience and personality is unrealistic. Thus, it is now claimed that breast versus bottle feeding by itself cannot ordinarily be sufficient to pattern the emotional aspects of personality. Rather, a series or pattern of events must be present, as for example, bottle feeding coupled with feeding on a fixed schedule, and with early and sudden weaning. In the present study, an attempt has been made to meet this condition by avoiding tests of relationship with an isolated infant training practice. Instead, the separate practices are combined in the form of indices containing several elements claimed to enter the pattern. These indices were constructed by assigning scores of one or more to each item making up the index. These scores or item weights were assigned on the basis of a review of the literature and inspection of the individual item distributions. None of the indices used have been subject to any more refined analysis or scaling technique, nor has any attempt at validation been made. Their sole claim to validity lies in their being composed of items derived from the literature on personality formation.

Similar indices were constructed in the case of the personality variables

derived from the Rorschach protocols and for similar reasons. In the case of the Rorschach test variables this procedure also provides certain safeguards which have long been advocated to eliminate possible biases and distortions in interpretation. Specifically, a charge frequently leveled against the Rorschach and other projective techniques is that the protocols can be interpreted to suit the preconceptions of the analyst. Therefore, in this study patterns of test responses have been defined *in advance*, and definite scores assigned to each. This procedure undoubtedly sacrifices some of the richness of less formalized methods of interpretation. But it was felt that such a procedure would provide the greatest accuracy and reliability.³

The children studied consisted of all third-grade children attending the village school in the highland village of Wattapola in March, April, and May of 1952 ($N = 34$); and a one-third interval sample of children attending the same grade of the Maradana Central School in the city of Colombo ($N = 39$). The Maradana school serves a working-class slum area. It is one of the better working-class areas of the city of Colombo and the socioeconomic status of the families studied was found to be roughly the same in Maradana as in the village of Wattapola. There was also no statistically reliable difference in the health of the rural and urban children as reported by their mothers.⁴

Interviewing and testing was conducted in Sinhalese by three women interviewers⁵ working under the author's direct supervision. Testing was done in the schools⁶ except in four cases of irregular attendance requiring testing in the home. The testing was, of course, done entirely in Sinhalese. The California Test of Personality was translated and reproduced in Sinhalese by mimeograph. The translation was done by the writer and two

³ Klopfer (7) scoring categories were employed. Full details of the construction of these indices are presented in Appendix 5 of the author's doctoral dissertation (15), together with the interpretation justifying the inclusion of the category in the index and the proportion of the sample giving the scored percepts. Scoring summaries of each protocol are also presented in Appendix 6.

⁴ The determination of equivalence in socioeconomic status and in health was made by use of indices designed to measure each of these variables. On the eleven-item socioeconomic status index the mean scores were 14.2 and 12.6 ($\chi^2 = 2.93$, $p > .20$). On the ten-item health index the mean score of the Wattapola sample was 4.4 and of the Maradana sample 4.6 ($\chi^2 = 0.14$, $p > .95$).

⁵ All three were senior students in sociology at the University of Ceylon and were familiar with the research problem. They received special training in the field techniques necessary for this study and worked under the writer's direct supervision. In Wattapola they lived in the village during the period of field work, as did the writer, but in Maradana they resided at home.

⁶ Strictly speaking, it was done during school hours. The physical location of the testing was in all cases outdoors, in order to let the children respond unobserved by their regular teachers and in order to get away from the noise and confusion which characterize Sinhalese schools.

students the previous year, checked by the chief interpreter of the Ministry of Education, and tried out in another village.

The conditions of measurement in this study make the problem of type II error an important consideration. The .05 confidence level would usually be set for a study of this type. However, in the present case the .05 level would mean a large risk of type II error. For example, a \bar{C} of as large as .47 based on a 3 by 3 table would have to be rejected as evidence of relationship in the universe when N is 65 and alpha is set at .05. Thus, in order to secure a better balance between the conflicting needs of minimizing the probability of not rejecting a false hypothesis or rejecting a true hypotheses, the .10 level has been chosen as best fitting the present situation.

Infant Feeding

Of all the infant experiences said to pattern the personality of an individual, feeding is usually considered of prime importance. The voluminous literature on the subject has been reviewed in several excellent published works. Honigmann (5) for example has succinctly summarized the purported relationships between infant feeding (and weaning) in Table 1.

For purposes of testing the hypothesis of relationship between oral gratification in infancy and personality, six practices were combined to form an Oral Frustration Index. The items comprising this index are shown in Table 2, together with weights assigned, and the proportion of the sample reporting each item.

TABLE 1
Some Early Experiences Important in Patterning Emotional Aspects of Personality after Honigmann (5)

Experience	Emotional Significance
Feeding (including sucking, nursing, regularity of feeding, nutritional aspects of food)	Earliest sense of security and adequacy is derived from the satisfaction of the alimentary needs; persistent nongratification interferes with the development of a sense of mastery, leading to pessimistic attitudes toward the world and the devaluation of the self. Unconscious equivalences with feeding may be set up which result in exaggerating the importance of oral and alimentary regions.
Weaning (abrupt, early, delayed, etc.)	Abrupt and early weaning may be traumatic; delayed weaning may become associated with emotional passivity. Unconscious equivalences with the mother as a frustrating object may appear in connection with abrupt or early weaning.

TABLE 2
Oral Frustration Index

Practice	Score	Percentage of Sample Reporting
Fed on schedule	1	53.7
Refusal to break schedule	2	38.2
Bottle-fed	1	41.8
Feeding of solids started at age:		
1-11 months	2	23.9
12-29 months	1	59.7
Weaning at under 24 months	1	19.4
Sudden weaning	2	26.9

The mean score on this index was 3.9 (S.D. 2.7) and scores occurred along the full range of possibilities from 0 to 9, thus providing a wide degree of variation on the basis of which variation in personality can be hypothesized to occur.

The personality test variables available from the California Test of Personality and the Rorschach technique are also well suited to the testing of the relationships postulated in Table 1. In Table 3, data on the relationship between the Oral Frustration Index and these personality variables is presented. Only two satisfy the .10 level of reliability.

Inspection of the mean scores for the three groups delimited on the basis of the Oral Frustration Index reveals a negative relation in the case of Inadequacy and a nonlinear but still somewhat negative relation in the case of Emotional Disturbance. That is, the more rigid the feeding practices which the child has experienced, the *less* likely is his Rorschach protocol to show signs of emotional disturbance or feelings of inadequacy.

It should be noted that in both cases the relationship between feeding practices and personality is the reverse of the hypothesis posed on the basis of the literature in this field. Moreover, no relationship was found between oral frustration and such other variables as the California "Sense of Personal Worth" test, or the Rorschach Insecurity rating, both of which should, on the basis of psychoanalytic theory, be associated with a rigid infant-feeding experience. In addition, the finding of only 2 relationships at the .10 level out of a total of 19 tests computed should be expected on the basis of chance alone provided there are no *a priori* grounds for those two specific relationships to be found in contrast to the other tests computed. The hypothesis of no relationship between infant-feeding practices and personality cannot therefore be rejected. It is concluded that the results of this investigation give no support to the view that infant-feeding practices are related to personality.

TABLE 3
Relation of Oral Frustration Index to Personality

Personality Measures	Average* Personality Test Scores for Children Whose Infant Feeding Experience Was:			χ^2	p	C
	Permissive (N = 25)	Intermediate (N = 26)	Rigid (N = 16)			
Rorschach Technique						
Maladjustment rating	12.0	13.2	11.6	1.79		
Emotional disturbance rating	2.0	2.5	1.9	10.79	.05	.504
Need for affection rating	2.1	2.1	1.7	2.78		
Anxiety rating	2.6	2.3	2.2	3.02		
Insecurity rating	2.2	1.9	1.5	6.79		
Aggression rating	2.2	2.2	2.2	4.24		
Passivity rating	2.1	2.0	1.7	2.97		
Inadequacy rating	2.7	2.0	1.9	7.97	.10	.442
Constriction rating	2.0	1.8	1.6	5.41		
California Test of Personality†						
Total adjustment	51.2	50.0	54.0	3.50		
Self-reliance	72.8	58.6	75.0	2.49		
Sense of personal worth	29.4	30.0	42.0	3.49		
Sense of personal freedom	61.0	28.0	26.7	6.39		
Feeling of belonging	41.6	46.7	45.0	3.14		
Freedom from withdrawing tendency	47.0	43.3	62.0	1.45		
Freedom from nervous symptoms	91.1	90.7	92.0	3.62		

* The mean is used in the case of the Rorschach data. For the California Test of Personality centile scores make the mean inappropriate and the median is used.

† The "Social Adjustment" subtests of the California Test of Personality are omitted from this table since there are no theoretical grounds for expecting them to be related to feeding. However, these tests were actually run, and no relationships beyond those expected by chance alone were found.

Elimination Training

A second basic Freudian character type, the anal erotic, is said to have its origin in the early sphincter training of the child. If such training is overly severe, compulsive and rigid attitudes toward many life situations develop. The data from both the present study and a previous study (14) show the sphincter training of Sinhalese children to be much less frustrating than the available data would indicate to be the case for European and American children. However, there is considerable variation. The index used to measure severity of elimination training has a maximum possible score of 11. The construction of this index is shown in Table 4. Scores occur over the full possible range, and the mean score is 6.4 (S.D. 2.2). On the basis of this variation, differences in the personality development may be expected.

TABLE 4
Anal Frustration Index

Practice	Score	Percentage of Sample Reporting
Age bowel training begun		
36 months and over	1	20.0
18-35 months	2	29.2
6-17 months	3	35.4
Bowel-training technique		
Mild ("asked him to go into jungle")	1	55.4
More severe (scolded, made him defecate after meals)	2	32.3
Scolded or punished in case of accident	2	6.2
Praised or rewarded for success	2	4.6
Age urinary training begun		
36 months and over	1	16.1
18-35 months	2	41.9
6-17 months	3	30.6
Urinary-training technique		
Mild (asked or told to use proper place)	1	73.4
More severe (scolded, taken to latrine regularly)	2	14.1
Night-training technique		
Mild (reminded him before bed)	1	26.6
More severe (scolded, taken to latrine regularly)	2	6.2

From Table 5 it can be seen that the Anal-Frustration Index used in this study is negatively related to the Rorschach ratings for Need For Affection and Anxiety. This indicates that the more severe the elimination training which the child has experienced, the less likely is he to give Rorschach percepts indicating excessive need for affection or anxiety.

These results are comparable to those obtained in testing the hypotheses concerning feeding practices. That is: (a) The finding of a negative relation is contrary to theoretical expectations. (b) No relation was observed between severity of elimination training and the personality variables which should theoretically be most highly related to it, namely the Rorschach scores for Feelings of Inadequacy and for Constriction. (c) A finding of only 2 statistically reliable relationships out of a total of 19 computed can be expected as a result of chance alone, especially when the observed relationships run counter to theoretical expectations. Thus, on the basis of the data presented in Table 5, the hypotheses of no relationship between severity of elimination training and personality cannot be rejected. It is concluded that the results of this study give no support to the view that sphincter-training practices pattern personality insofar as it has been possible to measure it in this study.

Before accepting the negative conclusions just presented, it is well to explore one further approach. It is generally argued that single Rorschach

TABLE 5
Relation of Anal-Frustration Index to Personality

Personality Measure	Average* Personality Test Scores for Children Whose Infant Elimination-Training Experience Was:			\bar{x}	p	\bar{c}
	Permissive (N = 26)	Intermediate (N = 23)	Rigid (N = 14)			
Rorschach Technique						
Maladjustment rating	13.2	12.6	12.9	5.10		
Emotional disturbance rating	2.2	2.4	2.0	2.47		
Need for affection rating	2.4	1.8	1.7	8.49	.10	.466
Anxiety rating	2.5	2.0	1.7	12.34	.02	.549
Insecurity rating	2.1	1.9	1.7	2.77		
Aggression rating	2.2	2.5	2.2	3.48		
Passivity rating	2.1	1.8	1.9	2.90		
Inadequacy rating	2.0	1.6	1.9	5.08		
Constriction rating	1.7	2.0	1.9	1.40		
California Test of Personality						
Total adjustment	56.7	52.1	60.0	6.77		
Self-reliance	72.5	60.6	73.3	1.20		
Sense of personal worth	43.3	29.2	28.0	2.84		
Sense of personal freedom	28.0	52.5	71.7	4.50		
Feeling of belonging	47.1	46.9	61.7	6.80		
Freedom from withdrawing tendency	46.2	47.0	45.0	.89		
Freedom from nervous symptoms	91.3	91.8	77.5	1.30		

* The mean is used for the Rorschach data. For the California Test of Personality centile scores make the mean inappropriate and the median is used.

scoring categories have little or no validity as personality indicators (1, 2, 7). It was for this reason that the technique of grouping several scoring categories together as indices of personality characteristics has been used in this study. However, it may be that these indices are faulty, and that this technique obscures rather than focuses the results. For this reason the relation between the Oral-Frustration Index and four standard Rorschach scores was tested. Elimination training was chosen for this methodological excursion since such theoretically expected personality characteristics as orderliness and concentration on detail are well represented by the Rorschach scoring categories.

A statistically reliable difference was found on only one of the four scoring categories shown in Table 6. Even that one provides no evidence to support the view that severe sphincter training is associated with the development of an anal-type personality structure since the most severely trained children in this sample had the lowest average rigidity of sequence score. Of the other (and not statistically reliable associations) two out of the three are also opposite to theoretical expectations.

TABLE 6
Relation of Anal Frustration Index to Selected Rorschach Categories

Rorschach Category	Measure	Infant Elimination-Training Experience			χ^2	<i>p</i>	<i>C</i>
		Permissive (N = 26)	Intermediate (N = 23)	Rigid (N = 14)			
Sequence	Mean*	1.7	2.5	1.6	9.11	.10	.446
F% over 50	%	80.8	82.6	85.7	.20		
D% over 70	%	42.3	52.2	35.7	1.23		
Dd% over 10	%	80.8	73.9	64.3	1.32		

* Obtained by scoring the usual sequence categories (orderly, rigid, loose, and confused) 3, 2, 1, and 0.

SUMMARY AND CONCLUSIONS

The wide variation in the infant-feeding and toilet-training practices which the sample children of this study have experienced provides an excellent basis for testing the hypothesis of association between these two aspects of infant care and personality. A total of 36 chi-square tests were computed for this purpose. None of the tests computed resulted in a statistically reliable association in the hypothesized direction.

The use of the California Test of Personality revealed no reliable associations in any direction. The use of the Rorschach technique resulted in five chi-squares with probabilities of .10 or less. However, all of these indicated a pattern of association opposite to that postulated on the basis of psychoanalytic theory. Specifically, the more indulgent and permissive the feeding experience of the children in the sample, the greater the tendency to give Rorschach percepts indicating feelings of inadequacy and emotional disturbance. Similarly, the more permissive and indulgent the elimination training of the children, the greater the tendency to respond with blot interpretations indicating anxiety and unmet need for affection.

In conclusion, it should be noted that these data do not require rejecting the importance of infant experiences for personality formation. But they do cast doubt on the view that any specific infant experience is related to any specific pattern of personality. At least three considerations give theoretical rationalization for the absence of such a specific pattern of association. First, the possible number of combinations of specific infant experiences is large. Until proved otherwise, it is soundest to assume that each of these combinations has a different effect. Second, the implication of the same experience may be expected to vary with the intensity of the experience. If this is true, the possible number of combinations is vastly increased, and the importance of any single experience still further reduced. Third, the meaning of any action varies from culture to culture. Short of evidence to the contrary, there is little reason for assuming that this is not

also the case in infancy. Thus, with three sets of variables operating, the possible number of combinations and permutations of specific infant experiences reaches an astronomical figure.

Nevertheless, these considerations do not mean that study of the personality significance of infant-training practices is without point. Nor do they deny the possibility that different societies may be characterized by a tendency for certain of the combinations to prevail. They do emphasize the often ignored need for precision in delimiting, examining, and generalizing from the available data. In any case, it may be concluded that the results of the present study give no support to the psychoanalytic formulation of the relation between infant-training experience and personality, and in particular the data fail to support the view that infant oral and anal gratification are associated with the development of either oral or anal personality structure.⁷

Manuscript received: August 20, 1956

Revised manuscript received: November 5, 1956

Murray A. Straus

Department of Rural Sociology

State College of Washington

Pullman, Washington

REFERENCES

1. Borgatta, E. F., and A. E. Eschenbach, "Factor Analysis of Rorschach Variables and Behavioral Observation," *Psychological Reports*, 1955, 1, 129-136.
2. Bosquet, H. T., and W. C. Stanley, "Discriminative Powers of Rorschach Determinants in Children Referred to a Child Guidance Clinic," *Journal of Consulting Psychology*, 1956, 20, 17-21.
3. Child, I. L., "Socialization," Chap. 18 in G. Lindsey, *Handbook of Social Psychology*, Cambridge, Mass.: Addison-Wesley, 1954.
4. Hallowell, A. I., "Culture, Personality, and Society," in A. L. Krober, *Anthropology Today*, Chicago: University of Chicago Press, 1953.
5. Honigsmann, J. J., *Culture and Ethos of Kaska Society*, New Haven: Yale University Press, 1949.
6. Honigsmann, J. J., *Culture and Personality*, New York: Harper, 1954.
7. Klopfer, B., M. D. Ainsworth, W. G. Klopfer, and R. R. Holt, *Developments in the Rorschach Technique: I.*, New York: World Book, 1954.
8. Lindesmith, A. R., and A. L. Strauss, "A Critique of Culture-personality Writings," *American Sociological Review*, 1950, 15, 587-600.

⁷ In evaluating these negative conclusions, the limitations of the data employed must be considered. Despite the confidence of clinicians in the Rorschach, its validity is by no means established even in Euro-American societies. The California Test of Personality items may be culturally bound and therefore not measure the equivalent characteristics in the society under study. Finally, despite efforts at standardization and control, the child-training data are also subject to error. Like most studies in this area, these data are based on recall, and all such data are subject to autistic and other distortions.

9. Orlansky, H., "Infant Care and Personality," *Psychological Bulletin*, 1949, 46, 1-48.
10. Ryan, B., and M. A. Straus, "The Integration of Sinhalese Society," *Research Studies of the State College of Washington*, 1954, 22, 179-227.
11. Sears, R. R., *Survey of Objective Studies of Psychoanalytic Concepts*, New York: Social Science Research Council, 1943.
12. Sears, R. R., J. W. M. Whiting, W. Nowles, and P. S. Sears, "Some Child Rearing Antecedents of Aggression and Dependency in Young Children," *Genetic Psychology Monographs*, 1953, 47, 137-236.
13. Sewell, W. H., "Infant Training and the Personality of the Child," *American Journal of Sociology*, 1952, 58, 150-159.
14. Straus, M. A., "Childhood Experience and Emotional Security in the Context of Sinhalese Social Organization," *Social Forces*, 1954, 33, 152-160.
15. Straus, M. A., *Child Training and Child Personality in a Rural and Urban Area of Ceylon*, University of Wisconsin, Ph.D. Thesis, 1956.
16. Thorpe, L. P., W. W. Clark, and E. W. Tiegs, *California Test of Personality, Primary Series*, Los Angeles, Calif.: California Test Bureau, 1942.

The Compensation of Executives¹

HERBERT A. SIMON, Graduate School of Industrial Administration

Carnegie Institute of Technology

In a recent analysis of the available data on the compensation of business executives, David R. Roberts (4) has shown that the compensation of the highest paid official in a company is related to size of company and to virtually no other variables (in particular, not to profit) after the effect of size has been partialled out. Specifically, the relation that Professor Roberts finds in his data is a logarithmic one. Let C be total annual compensation of the highest paid official, S , annual dollar sales, and a and k constants. Then the observed relation is:

$$[1] \quad C = kS^a$$

or, on a logarithmic scale:

$$[2] \quad \log C = a \log S + k'.$$

Fitting the data to equation [2] by the method of least squares, we find a value of about .37 for a . The data are homoscedastic on the logarithmic scale.

Professor Roberts discusses the implications of these data for the theory of executive compensation and advances an explanation based on marginal productivity theory. In the present paper I should like to develop an alternative theory of a more sociological character. This explanation has the advantage that, unlike the one based on the marginal productivity theory, it predicts not only a positive relation between size of company and compensation, but also the logarithmic form of the function and the approximate value of the coefficient a .

Businesses, like all large-scale organizations, are roughly pyramidal in form, because of the hierarchical structure induced by the authority relation. Each executive has a certain number, n , of subordinates at the level immediately below him, and this number varies within only moderate limits in a given company, and even among a number of companies. At executive levels it is seldom less than three, and seldom more than ten, and usually lies within narrower bounds—particularly if we take averages over all executives in an organization at a given level.

There is a widely accepted attitude in industry that an appropriate dif-

¹ This work has been carried out in the program of organizational studies at the Graduate School of Industrial Administration under a research grant from the Ford Foundation.

ferential in salary exists between an executive and his immediate subordinates, measured not in absolute terms but as a ratio. That is, an executive's salary "should" be b times the salary of his immediate subordinates, no matter what his level.² Again, the value of b undoubtedly varies from situation to situation, but one can find figures quoted in the range of 1.25 to 2. While we would expect to encounter instances of larger or smaller ratios, averages can be expected to be relatively stable.

Now, consider an idealized organization in which each executive has exactly n immediate subordinates, and in which he receives a salary b times the salary of his immediate subordinates. Let S be the number of executives in the organization³ and let L be the number of levels in the executive hierarchy. Then, we have the following relation between S and L :

$$[3] \quad S = 1 + n + n^2 + \cdots + n^{L-1} = \frac{n^L - 1}{n - 1} \approx \frac{n^L}{n - 1}$$

Now suppose that executives at the first, or lowest, level are brought in at a salary of A . Again, if this lowest level represents recent college graduates, there is good reason to suppose that there is a recruitment salary that does not vary widely from one position or one company to another. We then have the following equation to determine C , the salary of the top executive:

$$[4] \quad C = Ab^{L-1} = Bb^L$$

Taking logarithms in [3], we get:

$$[5] \quad \log S = L \log n + \text{const.}$$

Similarly, from [4] we get:

$$[6] \quad \log C = L \log b + \text{const.}$$

Eliminating L between [5] and [6], we find:

$$[7] \quad \log C = \frac{\log b}{\log n} \log S + \text{const.}$$

But equation [7] becomes identical with equation [2] if we take:

$$[8] \quad a = (\log b)/(\log n)$$

² This "rule of proportionality" receives prominent attention in most discussions of executive compensation, and its correctness as a norm is accepted more or less as a truism. See for example reference (3), and reference (2, pp. 319-321).

³ In the data we have available, the size of companies is measured by their dollar sales. We assume here an exact proportionality between dollar sales and total number of executives—which, of course, will hold only in an average sense. For this reason, we use the variable S indiscriminately to refer to both measures of company size.

We can test equation [8] further by seeing whether the observed value, .37, for a is consistent with reasonable values of b and n . Equation [8] defines b as a function of n , and vice versa, so that a whole set of possible pairs of values of the latter two variables will be compatible with a given value of a . We could have, for example, $n = 7, b = 2, n = 5, b = 1.75$, or $n = 3, b = 1.5$. All these pairs lie within the range we have postulated.⁴

There is one small additional piece of confirming evidence for our hypothesis. In 1936, General Motors made available data on the number of executives at various compensation levels. H. T. Davis (1), p. 49, found that this distribution could be described by the equation:

$$[9] \quad C' = mN^{-.33}$$

where N is the number of executives receiving compensation C' , and where m is a constant.

Under our previous assumptions, the number of persons at L' levels from the top is:

$$[10] \quad N(L') = n^{L'-1}$$

From the proportionality assumption, the compensation, C' of persons at this level is:

$$[11] \quad C'(L') = Mb^{1-L'}$$

Taking logarithms, and eliminating L' between [10] and [11], we get:

$$[12] \quad \log C' = -\frac{\log b}{\log n} \log N + \text{const.}$$

which becomes identical with the logarithm of [9] if we set $(\log b)/(\log n) = a = .33$. This new estimate of a is not very different from that obtained from the regression of salaries on size of company.

SUMMARY

In summary, I have proposed a theory of executive compensation that assumes that salaries are determined by requirements of internal "consistency" of the salary scale with the formal organization and by norms of proportionality between salaries of executives and their subordinates. Three mechanisms are postulated: (a) economic determination, through competition, of the salaries at the lowest executive levels where new em-

⁴ It is worth observing that if we try to estimate n and b directly in a particular firm, we will find the notion of "level" somewhat ambiguous, and will have difficulty distinguishing "levels" from "half-levels." However, any errors we make in estimating n will lead to proportional errors in our estimates of b , so that the ratio of $\log b$ to $\log n$ will be only slightly affected. For this reason, the test of our theory does not depend in any critical way upon the definition of level we employ.

ployees are hired from outside the organization; (b) social determination of a norm for the "steepness" of organizational hierarchies (usually called the span of control); and (c) social determination of a norm for the ratio of an executive's salary to the salaries of his immediate subordinates. Where these mechanisms operate, a relation will exist between the salaries of top executives and the sizes of their companies that matches very well with the observed relation. Moreover, if we take values that appear reasonable for the norms (b) and (c), we obtain a prediction of the slope of the relation between salary and company size that is in good quantitative agreement with the empirical data. Further, the same parameter values are obtained from data on the frequency distribution of executive salaries in a single company.

If the proposed theory is correct, it calls into question the usual economic explanation of compensation—that the executive is paid at a rate roughly equal to his marginal contribution to company profits. While the present theory is consistent with a positive correlation between compensation and ability, only an improbable coincidence would bring about equality between salaries determined by the mechanism described here and salaries determined by the marginal productivity mechanism. Hence, it would appear that the distribution of executive salaries is not unambiguously determined by economic forces, but is subject to modification through social processes that determine the relevant norms.

Manuscript received: August 1, 1956

Revised manuscript received: August 20, 1956

Herbert A. Simon

Graduate School of Industrial Administration

Carnegie Institute of Technology

Pittsburgh 13, Pennsylvania

REFERENCES

1. Davis, H. T., *The Theory of Econometrics*, Bloomington, Ind.: The Principia Press, 1941.
2. Koontz, H., and C. O'Donnell, *Principles of Management*, New York: McGraw-Hill, 1955.
3. Patton, A., "Current Practices in Executive Compensation," *Harvard Business Review*, 1951, vol. 29, no. 1, pp. 56-64.
4. Roberts, D. R., "A General Theory of Executive Compensation Based on Statistically Tested Propositions," *Quarterly Journal of Economics*, 1956, 20, 270-294.

Leadership Stability and Social Change: An Experiment with Small Groups¹

ELIHU KATZ, PETER M. BLAU, MORTON L. BROWN,
FRED L. STRODTBECK

University of Chicago

Recent reviews of research on leadership tend to discredit the idea that there are general leadership traits (5). The conception of "great men" (2) who inevitably emerge as leaders, no matter what the circumstances, competes with several alternative theories. It is frequently suggested that leadership is responsive to changing situations, and an individual who is a leader in one situation may not retain his position as the group moves on to another situation.

This need not mean, of course, that no distinctive mark sets off leaders from followers. For one thing, it may be that the distinctiveness of leaders does not rest on their attributes as such, but on the *relationships* between their attributes and those of the rest of the group. Homans (6), for example, makes this point when he concludes that the leader most closely embodies the norms and values of the group, and so, implicitly, do Merri (12) and Newcomb (14) when they show that leadership is contingent on the individual's willingness and ability to adopt the "traditions" of the group. Or, as Sanford suggests, "Leaders do have traits which set them apart from followers, but *what* traits set *what* leaders apart from *what* followers will vary from situation to situation" (17, p. 51).

The concept of situation has many dimensions, and it is not clear which of these are responsible for leadership stability and change. One specific dimension of a situation is its "content"—that is, the particular *activity* in

¹ This experiment was carried out at the University of Chicago within the framework of two courses in small-group analysis offered jointly by Strodtbeck, Blau, and Katz. Brown, a member of the original classes, served as project assistant during the analysis of the data and completion of this report. Other class members to whom we are indebted include Philip Cohen, Athol Congalton (Visiting Rockefeller Fellow from New Zealand) Charles Hawkins, Nathan Kantrowitz, Joe Spaeth, Rodney White, Kathleen Wiseman, and in particular, Frank Cherry, who did the interaction process scoring, and John Johnstone, who completed the preliminary tabulations. Messrs. Ernest Kahn, Steve Katz, and Jack Wank of the Jewish Community Centers of Chicago, and Mr. Elliot Eisner, of the South Shore Temple, facilitated recruitment of subjects. Financial assistance from the Social Science Research Committee of the University of Chicago and a Ford Foundation grant-in-aid to one of the authors are gratefully acknowledged, as is permission to use facilities of the Law School Jury Project and of the University Radio Office.

which the members of the group are engaged. In fact, the situational approach to leadership appears to imply that leaders rise and fall according to the changes of activity in which their groups are involved.

It cannot be so simple, however, for groups are constantly changing their activities. The interaction of men on the job, for example, leads to the formation of on-the-job-friendships which, in turn, leads to a succession of new activities such as helping one another at work, eating lunch together, playing games, and so on. In the normal course of any group's development, Homans suggests, the range of activities in which the group engages will be continually elaborated (6, pp. 109-110).

But if the "range of activities" can be taken to mean the number of new activities in which the group engages (helping, lunching, playing games), the question arises whether group leaders also rise and fall in response to this sort of situational change. An unrefined interpretation of the "situational approach" would suggest that they do; intuitively, one would expect leaders to be more stable than that. For the purpose of the present research, we may ask more generally: *Under what conditions are changing activities associated with changes of leadership, and under what conditions does the leader survive from one situation to the next?*

One clew for answering this question comes from the work of Carter and Nixon (3). These experimenters find that changes from an intellectual to a clerical task do not produce changes in group leadership, whereas shifts from intellectual to mechanical tasks do. This suggests that leaders change not with every change of activity, but only with changes to markedly different *kinds* of activity—that is, to activities which may no longer be compatible with the leader's talents.

Searching the literature for other examples of leadership change takes us far beyond the small group laboratory. Studies of disaster have pointed out that leadership and initiative in such emergencies arise from unexpected quarters (8). Leaders in Officer Candidate School often fail to retain their leadership in combat (17). The authority of the village patriarch has been observed to falter in situations of urbanization and immigration (4, 10, 11). Similar observations have been made in connection with programs of technical assistance to underdeveloped areas, where concern has been expressed over the consequences which arise for the social structure as a result of the introduction of technological change (15). Despite their variety, these examples have a factor in common. In all of them the impetus to change originated *outside* the group.

But unlike disasters or technical assistance programs, the majority of changes that take place in the activities of most groups are everyday changes that occur at the group's own initiative. Here, perhaps, is a second important clew to the conditions of leadership stability and change: if a

new situation is beyond the group's control to accept or reject, as in a disaster, the leader's position will be jeopardized.

This implies that if the group does have a choice in deciding on the activities in which it is to become involved, the leader will have a better chance of retaining his position. One reason for this might be that, in such a case, the leader will try to influence the group's choice of activities in accordance with his own interests and capabilities and, since he is the leader, he may be expected to have a good chance of succeeding in this attempt. In other words, leaders may secure their positions by promoting only those new activities which they perceive to be conducive to their continued leadership.

This paper reports the results of a study of the effects on leadership of various changing conditions—both externally imposed changes and those that arise as the result of the group's own social development. To permit the investigation of the emergence and subsequent fate of leadership, and of the rise and social development of a group, *ad hoc* groups of high-school students were assembled. Each group performed a succession of three tasks undertaken under different experimental conditions and, fundamentally, the experiment concerned the extent of leadership stability and change under these varying circumstances. Before considering the guiding hypothesis more explicitly, however, we shall report on the experimental procedures employed.

THE EXPERIMENTAL PROCEDURE

Time One (Original Discussion): A group of three or four high-school students was seated around a small, square table at positions labeled A, B, C, and D. The experimenter told the group that a study had found that popular teen-age girls tend to prefer "blue" rather than "happy" songs. The group was asked to discuss possible explanations of this finding and, in ten minutes, to arrive at a unanimous answer, if possible. The experimenter then seated himself at a table in the far corner of the room with his back to the group. Unless he was summoned earlier (which was infrequent), he warned the group after eight minutes that it had only two minutes left, and interrupted the discussion after two minutes had elapsed to ask for the group's opinion.

During this discussion and throughout the experiment, two observers sat behind a glass partition. They could be seen dimly, but this did not prove to be distracting. One observer scored interaction according to Bales' system (1); the other operated the tape-recording equipment and observed manifest acts of leadership.

Following the first task, subjects were asked to rank members of their group, including themselves, in terms of the question, "Who do you think

contributed most to the solution of the problem?" They were also asked to rank their fellow group members, not including themselves, on the basis of the question, "Whom did you enjoy working with most?" Both rankings were performed by ordering sets of lettered cards, the letters corresponding to the signs in front of the group members.

From this first task at Time One, a leader can be located. In this paper, our attention will be restricted to the measure of leadership based on the "contributed most" post-task rating.

Time Two (Selection of a New Task): Following the collection of the rankings, in the period designated as Time Two, the experimenter briefly described to the group three new tasks: (a) another discussion—namely, acting as a mock jury to decide whether or not the sale of horror comics should be prohibited; (b) a paper-and-scissors puzzle, requiring group collaboration; and (c) a construction task, requiring the group to build a house or some other object out of plastic panels. Each member was asked to rank the three tasks on a secret ballot according to his relative preference for each as the group's next task. When the ballots were collected, the members of the group were asked to discuss together what to do next, and to come to a common decision.

From Time Two, the experimenter obtains on the ballot the original task preference of the leader and of other members of the group and, following the deliberation, the ultimate decision of the entire group. Thus, in Time Two we can examine the leader's personal preference and the extent of his influence over the group discussion and decision.

Time Three (Chosen or Imposed Task): Half the groups (every other one) were permitted to carry out the task they had chosen. The other half were told by the experimenter to do one of the tasks they had not chosen. Thus, at Time Three, groups performed a task which was either chosen or imposed. As in Time One, the duration of the task was approximately ten minutes. Interaction was again recorded, and the task was followed by the administration of the two sociometric tests.

The differences in the sociometric rankings between Time One and Time Three provide a criterion for comparing leadership stability as the following conditions vary: 1. The leader's preference did or did not prevail over the others in the Time Two decision; 2. The Time Three task was one *chosen* by the group or *imposed* by the experimenter; and 3. The Time Three task was relatively analogous to the Time One discussion (the jury task), or it was less analogous (the puzzle or construction task).

Time Four (Assignment of Alternative Task): Those groups which performed the chosen task at Time Three were assigned a new task (nec-

essarily, one they had not chosen), while those groups which had had a task imposed upon them at Time Three were permitted to do what they had requested at Time Two. Again, interaction was recorded and sociometric questions were administered.

The results of Time Four permit us to ask again the same kinds of questions asked at Time Three as well as additional ones concerning the effect of the group's development over another period of time.

Subjects were all teen-agers who belonged to youth groups at four community centers on the South Side of Chicago. Within an experimental group subjects were of the same sex and age. In general, subjects were known to each other only slightly or not at all before arriving in the test room. Twenty-five of the 39 groups consisted of 4 members each; 14, because of absenteeism, had only 3 members. The entire experiment was performed at one sitting, taking about an hour.

The Guiding Hypothesis

Our central hypothesis is that leadership will be more likely to remain stable when a group has its own way about "elaborating the range of its activities" than if new demands are thrust upon the group from outside (cf. 7, p. 129, fn. 16). Underlying this hypothesis are several assumptions about the process at work: (a) First, it is believed that leaders will want to retain their positions of leadership. To accomplish this (b) leaders will use their influence to guide their groups into those activities which they perceive as hospitable to their continued leadership. Therefore, it is also anticipated that (c) when given the choice, leaders will recognize those tasks in which they can continue to lead. These are likely to be tasks which (d) are "analogous" to the ones in terms of which they rose to power in the first place.

Testing Several Assumptions

Turning to the results of the experiment, we shall first test some of the assumptions proposed here about the dynamic process of maintaining leadership. Let us begin at the end of our list of assumptions by asking whether the leader was disproportionately attracted to the jury task—that is, to the task most obviously analogous to the discussion task in terms of which he rose to leadership. Our findings indicate that this was not the case. Men of all ranks favored the jury task to the same degree. Contrary to our expectation, that is, individuals who had risen to leadership in the course of a discussion task did not express their preference for another discussion task more often than did nonleaders. The jury task was an overwhelming favorite for all ranks.

Rank at Time One	Per Cent and Number Preferring Jury Task
1	64 (39)
2	54 (39)
3	64 (39)
4	68 (25)*

* (There were 25 four-man groups and 14 three-man groups.)

A second bit of evidence mildly supports the findings previously reported by Carter and Nixon (3). When the analogous jury task *was* actually performed, 5 out of 10 groups retained their Time One leaders, and when the nonanalogous tasks were performed, only 11 out of 29 groups retained their leaders. This difference, however, is not significant and, even with further cross-classification, the analogous vs. nonanalogous distinction does not become discriminative. It should be noted that one possible reason for this negative finding is that the discussion of song preferences of girls and the jury discussion were actually not so analogous as had been assumed, but required quite different skills.

A third piece of evidence is more important, however. Whether or not the leader chooses an analogous task, and whether or not the analogous task is actually performed are both less relevant for leadership stability than whether or not the leader performs the task of his choice—whatever the content of that task happens to be. Six of the ten leaders who performed tasks for which they had expressed preference maintained their leadership compared with only 10 of the 29 leaders who performed a task which they

Leader's Preferred Task	Per Cent of Groups with Same Leaders at Time One and Time Three
Performed	60 (10)
Not performed	34 (29)

did not choose. This difference, almost significant at the .05 level, suggests that leaders either do recognize the tasks in which they are most apt to maintain their leadership, or that leaders who made the choice that subsequently proves to have been the "right" one are more likely to retain the allegiance of the rest of the group.

But our data also show that Time One leaders were no more able than others to influence their groups to choose the tasks which they themselves preferred. While it still may be that this assumption is justified for more permanent groups, in our groups, at any rate, the sociometrically defined leader's influence on the group decision at Time Two was not greater than that of any other member. If we use the term *prevailed* to indicate simply

the correspondence of a personal decision with the ultimate group decision at Time Two, then it is apparent that no consistent difference characterizes the several ranks.

Rank at Time One	Per Cent Prevailed
1	41 (39)
2	41 (39)
3	51 (39)
4	44 (25)

The Experimental Results

Having in this way disposed of several of the underlying assumptions, we shall now turn to consider the effects of permitting the group to "choose" its activity in contrast to "imposing" an activity. Disregarding for the moment whether our assumptions concerning the underlying process are warranted, we still want to know whether our central hypothesis is tenable: Do new activities which are chosen by the group in the course of its own interaction tend to conserve the existent social structure better than activities which are imposed by an external agent?

Treatment	Per Cent of Groups with Same Leaders at Time One and Time Three
Chosen task	50 (20)
Imposed task	32 (19)

That leadership stability is slightly higher in the groups which performed tasks which they had chosen is consistent with the original hypothesis, but the magnitude of the difference is disappointingly small. Yet if we examine this variable while holding constant whether or not the leader prevailed in the group decision at Time Two, we find that the *combined effects* of both factors does make a significant difference. Consider first the left-hand side of Table 1. If the leader's personal opinion prevailed in the

TABLE 1
Leadership Stability between Time One and Time Three for Chosen and Imposed Tasks according to Whether the Leader Prevailed

Leadership Stability	Leader Prevailed		Group Unanimous		Leader Did Not Prevail	
	Task Chosen	Task Imposed	Task Chosen	Task Imposed	Task Chosen	Task Imposed
Stable	4	1	2	2	4	3
Unstable	2	9	2	2	6	2
Total	6	10	4	4	10	5

discussion his chances of remaining leader appear to be contingent on whether or not the group is allowed to work on the task of its choice. This is in accord with the hypothesis. In nine out of the ten cases where the task was imposed, the leader had to surrender his position, while he lost his leadership in only two of the six cases where the group carried out its chosen task. (This difference is significant at the .05 level.)

If the leader did not prevail in the discussion, however, the relationship between task assignment and leadership stability is quite different. Here, the leader has not succeeded in persuading the group to go along with his preference. When this happens, the chances of the leader retaining his position were three out of five if the task was imposed and only two out of five (i.e., four of ten) if the task were chosen. It might be supposed that imposing a task on a group that rejected its leader's preference is simply reinstating the task that the leader had originally desired; this is not the case, however, since—by a sheer coincidence—the task imposed was in every case the task that was neither the leader's private preference nor the one the group had decided upon. Thus, on the right side of Table 1, the relationship between task assignment and leadership stability actually becomes reversed: the imposed task produces greater stability than the chosen task, and neither is a task which the leader desired.²

A crucial test of the impact of the experimental intervention of chosen and imposed tasks under the conditions where the leader did or did not prevail is to be found in the second-order interaction between the respective sides of the table. If the effects were the same for both parts, then the ratio $\frac{ad}{bc} / \frac{a'd'}{b'c'}$ would be 1.0, the letters referring to the values of the respective cells of Table 1. It may be observed that the ratio is in fact very much higher: $\frac{4 \times 9}{2 \times 1} / \frac{4 \times 2}{6 \times 3} = 18/.44 = 41$, which exceeds the jointly estimated expected value by an amount which is significant beyond the .01 level (as determined by a chi square test with 1 degree of freedom). It is also useful to note that when the group decision at Time Two was unanimous (the center of Table 1), the four chosen and four imposed cases result in a degree of stability which is intermediate to that found at the extremes.

In sum, the effect on leadership of the experimental assignment of task depends on the process by which the group has reached its decision. If the group's choice of task was guided by the leader, the imposition of a different task threatens his leadership, as originally predicted. If, however, the group was not guided by the leader in its choice of task, the imposition of

² To simplify exposition, the complex analysis of the effects of the preferences of all groups members on status is not reported here. Let it suffice to say that this factor does not account for the negative cases in Table 1.

a task other than the one preferred by the group does not threaten his position. On the contrary, this double negative—the experimenter's rejection of the group's preference which itself had been a rejection by the group of the leader's preference—seems even to strengthen slightly the leader's position.

A Nonexperimental Source of Variability

Experiments with *ad hoc* groups, as compared with the study of actual, ongoing groups, have both the advantage and disadvantage of magnifying the effects of experimentally introduced variations. In real life, even relatively impermanent groups, in the very process of being formed, develop a social structure that influences their reactions to changes in external conditions. Since the groups with which we worked had been in existence barely twenty minutes when the experimental variable (chosen vs. imposed tasks) was introduced, one may suspect that the results obtained might have been due entirely to the vulnerability of a twenty-minutes-old social structure. In actuality, however, minute as the historical development was, internal processes beyond the control of the experimenter had already given rise to structural elements which exerted a pronounced influence on leadership stability. In other words, even an experimental group has a "history," and this history may not only have consequences for whatever is being studied but may also obscure the effects of the experimental operations.

The interaction process scores (1) which preceded the Time One rankings reveal that there are differences governing the patterns of relations among group members in various groups from the very moment of first meeting. A dimension which seems to be particularly sensitive as an early indication of the group's developing structure is the amount of disagreement directed to the incipient leader during the discussion at Time One. Percentage of negative acts received (categories 10, 11, and 12) by the individual who is later to be named as leader is used as the measure of disagreement. Dividing leaders at the median (where negative acts are 16 per cent of all acts received), it may be seen in Table 2 that those who received a higher proportion of negative acts were less likely to retain their leadership than the others. For this table, chi square is significant at the .05 level.

From the standpoint of the experimental design, this is indeed a sobering insight because it implies that the leader, in the course of his interaction at Time One, has sown the seeds of his own downfall before the experimental treatment has been applied. It should be noted that the extensive disagreement did not prevent a rise to leadership at Time One, but, as has just been demonstrated, it is related to leadership stability—that is, to whether the Time One leader remains leader at Time Three. This suggests that an individual can become the leader in a small group despite behavior

TABLE 2

Leadership Stability and Extent of Disagreement with Leader at Time One

Leadership Stability	Extent of "Disagreement" with Leader		
	Low	High	Total
Stable	12	4	16
Unstable	9	14	23
Total	21	18	39

that evokes antagonism and disagreement from other members, but that leadership achieved in such fashion is less stable than leadership won without antagonizing the others.³

The relation between disagreement received and stability is, of course, not presented as exhaustive of the sources of uncontrolled variability. It is rather regarded as a particularly clear example of the way in which internally developing processes in a group become implicated in experimental designs and in experimental results. Since social processes spontaneously generated in groups influence leadership stability, it is desirable to understand how these processes condition the effects of the experimental variables. Let us, therefore, examine the effect on leadership stability of a combination of the internal process of "disagreement" and the two variables we have previously examined: (a) whether the leader prevailed or did not, and (b) whether the task was chosen or imposed.

Each cell of Table 3 is an expression of the extent of leadership stability under a combination of the three conditions. The frequencies, of course, are very small but the principal objective—to show that the information in Table 1 is not confounded with that in Table 2 and that the relationship reported in each is independent of the other—is achieved. It is to be noted that the previously observed interaction effect between prevailing in the discussion and whether the task is chosen or imposed (Table 1) is clearly evident in Table 3 *provided that* the leader has experienced little disagreement in his rise to leadership (upper row). Under these conditions, leaders were stable in all groups that agreed to the leader's preferred task and then carried it out, as well as in all groups that rejected the leader's preference only to be rejected, in turn, by the experimenter. In contrast, only about a third of the leaders in other groups maintained their positions.

³ It is interesting to observe that Time One disagreement is more predictive of Time Three leadership stability than is disagreement at Time Three. Landsberger (9, p. 566, Table III) shows that only early disagreement is related to satisfaction with the outcome of a labor-management mediation session. Mills (13) also reports that only early, but not late, disagreement, has important consequences for interpersonal evaluation.

TABLE 3

Percentage of Groups with Stable Leaders as Experimental Conditions as Well as Internal Group Developments Vary

Extent of "Disagreement" with Leader	Leader Prevails or Group Unanimous		Leader Does Not Prevail	
	Chosen	Imposed	Chosen	Imposed
Low	100 (4)	37 (8)	33 (6)	100 (3)
High	33 (6)	0 (6)	50 (4)	0 (2)

However, when there was a high level of disagreement with the leader in the formative period of the group (bottom row), there is a somewhat different story to tell. Here, first of all, there is a generally lower level of stability; in three of the four cells, leaders who experienced high disagreement are less stable than those who experienced low disagreement. Second, high disagreement seems to nullify the effect of whether or not the leader prevailed. Both when leaders prevailed and when they did not, no leader survived the impact of an imposed task.

A Further Test of the Situational Hypothesis

While the finding that the effects of experimentally introduced variations were modified by uncontrolled aspects of the group structure compromises our claim to a rigorous experiment, it also indicates that it is possible to reproduce functioning group structures in a few minutes in the laboratory. If something of the bold orientation to the present which characterizes so much of small group research has been sacrificed, perhaps some advance has been made toward an understanding of the interplay between externally controlled influences stemming from outside the group and the processes which characterize the normal course of group development.

It appears that the imposition of an unwanted task threatens leadership stability particularly when the leader has already undermined his position by generating opposition and dissatisfaction, and when, although he experienced low disagreement, the task the group asks for but does not get is one that the leader has advocated. Strangely enough, leaders who experienced little opposition maintained their positions not only when the tasks which they preferred were seconded by their groups and then carried out; they were equally stable where their choices were rejected if the group choice was vetoed by the experimenter.

It will be recalled that at Time Two the group could select from among three alternative tasks. Had there been only two such tasks, the effect of imposing a task upon a group which had rejected the task preferred by its leader would have been to force the group to carry out its leader's prefer-

ence. It might have worked out this way even when there were three tasks, but, in fact, it did not. As pointed out earlier, it so happens that the task imposed in all cases where the leader did not prevail in the Time Two decision was not the one the leader had chosen, but the third alternative which neither leader nor group majority preferred. Leadership stability in such cases cannot be attributed, consequently, to the fact that groups simply worked on tasks which their leaders preferred. It seems rather that in overruling the group which had rejected its leader's choice, the adult experimenter had weakened opposition to the leader and thereby indirectly reinforced his position.

This interpretation of an unexpected finding calls into question the major hypothesis that, given certain conditions, leadership stability depends on the group's freedom to choose the situation in which it becomes involved. For if the experimenter's intervention, and not the experimental variable itself, explains leadership stability in those cases where the leader did not prevail, it may also be the factor responsible for stability in those where he did. When the leader has prevailed in the discussion, that is, and the group is then assigned the task it has chosen, this action of the experimenter may have been interpreted as implicit approval for the leader's guidance. It is possible that this reinforcement of the leader's position, rather than obtaining the preferred task, is responsible for the high stability of leadership observed under these conditions. In short, we are asking whether experimentally produced effects, just as those in Roethlisberger and Dickson's famous illumination experiment (16), resulted from the substance of the experimental variations or from the process of introducing them?

This question can be answered by deriving different predictions from these two possible interpretations. Let us concern ourselves with the ten groups which accepted their leaders' preferences (or unanimously agreed) in the group discussion and who were accorded the task of their choice by the experimenter (columns 1 and 3 in Table 1 combined). Let us ask whether the relatively high stability observed in these groups is due to the fact that the leader can perform the task he desired or to the fact that the experimenter is reinforcing the leader's status by assigning the task that the leader preferred. By examining what happened at Time Four, it is possible to see which interpretation carries more weight.

At Time Four, it will be recalled, all groups were asked to perform a task other than the one performed at Time Three. Since the ten groups we are here discussing received a task of their own choosing at Time Three, at Time Four they necessarily were asked to perform a task they had explicitly rejected. If the "situational" hypothesis is correct—that leaders change in response to the changing activities which confront their groups—then these leaders ought to have only a very poor chance to retain their

leadership at Time Four. If, however, the "reinforcement" hypothesis is correct, the fact that these leaders had all been supported by the experimenter previously (in assigning the Time Three task) might be expected to stand them in good stead at Time Four as well.

The results are these: Whereas six of the ten Time One leaders remained leaders at Time Three, only one was again a leader at Time Four. The fate of these Time One leaders at Time Three was nearly twice as good as those of other Time One leaders (10 of 29 remained leaders), but their chance of retaining or resuming leadership at Time Four was less than one third as good as those of other Time One leaders (11 of 29 were leaders again at Time Four). Tentative as any conclusion based on so few cases must be, these data seem to indicate that the experimental variation of chosen or imposed tasks, and not merely the reinforcements incidental to the introduction of these variations, produced differences in leadership stability.

DISCUSSION AND SUMMARY

An experiment was conducted with 39 groups of 3 or 4 high-school students. The guiding hypothesis was that leadership will be more stable in groups working on tasks of their own choice than in those performing tasks imposed upon them.

Two kinds of variables have proved effective in the analysis of the results of this experiment: (a) those which are the outcomes of group process, for example the disagreements addressed to the leader at Time 1 and whether the leader prevailed at Time 2; and (b) those controlled by the experimenter—the chosen or imposed tasks. Our results suggest that the effects of the experimental variable itself are conditioned by its position in a sequence of "emergent" events—that is, whether the leader had received high or low disagreement and whether the group had seconded his preference. To a certain extent, then, we have found it necessary to sacrifice some amount of experimental rigor in order to take account of the specific sequence of events—that is, of "history."

The importance of spontaneous group processes (such as high-low disagreement) suggests that future research ought to make explicit provision to take account of them. To dispose of the difficulty by experimental elimination of the "emergent" group processes is no solution because it makes it impossible to study how group development conditions the effects of externally introduced changes. The remedy is not to ignore these processes but to design experiments that permit their systematic investigation.

It ought also to be noted that this experiment was not concerned at all with the initial determinants of Time-One status. Why some individuals emerged in rank 1, and others did not, is a legitimate concern of leadership study but one which we did not investigate.

As for our underlying assumptions concerning group process, we are reluctant to give them up: that leaders want to persist as leaders; that leaders will use their greater powers to direct the group to tasks compatible with their leadership; that tasks analogous to those in which leaders arose will be perceived by them as more compatible. We found that the leader's preference was not more often reflected in the group decision than that of other members, and we found that leaders did not choose the analogous task more often than did other members. At the same time, we found that whether or not the leader actually performed the task of his choice, analogous or not, he was more likely to be stable than if he performed a task which was not of his choosing. There was also a suggestion that performing the analogous task resulted in slightly greater stability than performing the nonanalogous task. These matters deserve further investigation and should not yet be considered resolved.

As for the central concerns of the paper, the following were the major findings:

1. Leaders who experienced much disagreement in the course of establishing themselves were subsequently more prone to lose their status than other leaders who aroused less antagonism. In the groups where leaders were strongly opposed at first, experimental variations produced the expected results: leaders of groups which received an imposed task all lost their leadership (whether or not their preference had been seconded by the group), while four of the ten in groups receiving chosen tasks were stable.

2. When the leader experienced little disagreement and his choice of task was also accepted by the group, the expected effect of the experimental variable was most pronounced. If the task was chosen, all leaders retained their leadership; if it was imposed, only a third was stable.

3. When the leader experienced little initial opposition but his choice of task was rejected by the group, contrary to expectations, he was more likely to remain a leader if the group had a task imposed upon it. Apparently, when a group that challenges its leader is overruled by the superior authority of the experimenter, its nascent opposition is undermined and the leader's position is strengthened. This raises the question of whether the process of introducing the experimental variation rather than the nature of the variation itself is responsible for leadership stability and change. The pattern of leadership at Time Four suggests that the experimental variation was the important factor.

Manuscript received: August 13, 1956

Revised manuscript received: October 29, 1956

*Fred L. Strodbeck
Graduate Program in Social Psychology
University of Michigan
Ann Arbor, Michigan*

REFERENCES

1. Bales, R. F., *Interaction Process Analysis*, Cambridge, Mass.: Addison-Wesley, 1950.
2. Borgatta, E. F., A. S. Couch, and R. F. Bales, "Some Findings Relevant to the Great Man Theory of Leadership," *American Sociological Review*, 1954, 19, 755-759.
3. Carter, L. F., and M. Nixon, "An Investigation of the Relationship between Four Criteria of Leadership Ability for Three Different Tasks," *Journal of Psychology*, 1949, 27, 245-261.
4. Eisenstadt, S. N., "Communication Processes among Immigrants to Israel," *Public Opinion Quarterly*, 1952, 16, 42-58.
5. Gibb, C. A., "Leadership," in G. B. Lindzey, *Handbook of Social Psychology*, Vol. II, Cambridge, Mass.: Addison Wesley, 1954.
6. Homans, G. C., *The Human Group*, New York: Harcourt Brace, 1950.
7. Katz, E., and P. F. Lazarsfeld, *Personal Influence*, Glencoe, Ill.: Free Press, 1955.
8. Killian, L. M., "Some Accomplishments and Some Needs in Disaster Study," *Journal of Social Issues*, 1954, 10, 66-72.
9. Landsberger, H., "Interaction Process Analysis and the Mediation of Labor-Management Dispute," *Journal of Abnormal and Social Psychology*, 1955, 552.
10. Lerner, D., P. Berkman, and L. Pevsner, *Modernizing the Middle East*, forthcoming; see also D. Lerner, "The Grocer and the Chief," *Harper's*, Sept. 1955.
11. McPhee, W. N., and R. B. Meyersohn, *The Radio Audiences of Lebanon*, unpublished research report of the Bureau of Applied Social Research, Columbia University.
12. Merei, F., "Group Leadership and Institutionalization," in G. E. Swanson, T. M. Newcomb, and E. L. Hartley (eds), *Readings in Social Psychology*, New York: Holt, 1952.
13. Mills, T. M., "Power Relations in Three-Person Groups," *American Sociological Review*, 1953, 18, 351-357.
14. Newcomb, T. M., "Attitude Development as a Function of Reference Groups: The Bennington Study" in G. E. Swanson, T. M. Newcomb, and E. L. Hartley (eds), *Readings in Social Psychology*, New York: Holt, 1952.
15. Paul, B. D. (ed.), *Health, Culture, and Community*, New York: Russell Sage Foundation, 1955. (See articles by O. Lewis, "Medicine and Politics in a Mexican Village," 403-504, and K. Oberg and J. A. Rios, "A Community Improvement Project in Brazil," 349-376.)
16. Roethlisberger, F. J., and W. J. Dickson, *Management and the Worker*, Cambridge, Mass.: Harvard University Press, 1939.
17. Sanford, F. H., "Research on Military Leadership," in J. C. Flanagan (ed.), *Psychology in the World Emergency*, Pittsburgh: University of Pittsburgh Press, 1952.

Political Opinion and Personal Security¹

GWYNNE NETTLER, *Monterey Peninsula College*

JAMES R. HUFFMAN, *United States Air Force*

"Reformers blame the world for being themselves ill fitted to live in it" (37).

Santayana is not alone in his assumption that social criticism represents an ideological transvaluation of personal discomfiture with others. Writers of diverse political hues have penned their suspicions that radicalism is born of resentment and social "dis-ease." Thus Weber (52), Nietzsche (33), Scheler (38), and Toynbee (50); Wolfe (54), Schumpeter (39), and Lynd (25); Janet (18), Mayo (28), Plant (34), and E. Jones (19); Lasswell (23) and Terman (47), and even the "radical" Norman Thomas himself (48).

The list of theoretical contributors is impressive; empirical substantiation less so for the usual reasons of defect of sample or of instrument, or deviousness of approach.

Some empirical support of the thesis equating personal disgruntlement with political discontent is given by studies of Freudian implication which find modest, but significant, associations between resentment of parents and resentment toward society (3, 22, 30, 44).

Secondary evidence of this relationship is derived from the reported positive correlation of social status and personal adjustment (4). Since the positive relationship between socioeconomic status and conservatism is even more strongly documented (7), one might deduce a similar association between personal adjustment and conservatism.

Direct correlational studies of political opinion and personality report a consistent, low association of introversion with radicalism. Dexter (8) finds a correlation of $+ .24$ for her women students, Sanai and Pickard (36) report an insignificant coefficient of $+ .16$ for 70 Englishwomen, and Gundlach (14) and Howells (17) report similar low-order correlations. Bray (5) finds an association of "nervous tenseness" with nonconformity which may be relevant to our problem. Krout and Stagner (22) differ from these investigators in finding no association between introversion and radicalism, although they do find their radicals to be "socially inadequate."

Diamond (9) holds that introversion is maladaptive and associated with informal or nonassociational radicalism. He goes on to show how the assumption of radical ideology and absorption in radical-party activity changes personality in the direction of extroversion and greater social ease.

¹ Funds awarded by the Research Committee of the University of California, Santa Barbara College, substantially assisted the conduct of this study and acknowledgment is hereby gratefully accorded.

It seems fair to conclude from these studies that an introverted bookishness, withdrawal from others, friction in social relations, nervous tenseness—in other words, social “dis-ease”—is more characteristic of the individuated, degrouped radical than of the conservative or the group-immersed radical. This conclusion needs qualification, of course, in terms of the predominantly collegiate composition of most samples used.

However, in interpreting the persistent, low-order association of introversion with radicalism, some investigators deny any taint of “maladjustment” or “neurosis” in such inward-orientation. Howells, for example, calls the introversion of his “religious liberals” nonneurotic (17), and Newcomb (30) holds that “none of the neurotic aspects of introversion . . . appear to characterize the liberals.”

Of course, interpretation here depends upon what one means by “neurotic” and “introverted.” Certainly the Jungian conception of introversion implies lesser adjustment (21).

THE INSTRUMENTS

The confusion as to what is being correlated with political opinion—whether it be “personal adjustment” or some quality of this vague label such as introversion or neuroticism—this confusion, plus the questionable validity of some measures of political opinion, have motivated our study.

The S-I Scale

Certainly the concept of “personal adjustment” is a broad and vague one, with even less definition than “radicalism-conservatism,” and the labels “adjusted” and “integrated” are probably more expressive of the investigator’s preference than they are descriptive of personal traits.

At first it was thought that the area of interpersonal relations relevant to our problem could be sufficiently narrowed and clarified by selecting “neurosis” as the behavior variable and some standardized “neurotic inventory” as the measure. However, perusal of the literature on neurotic behavior indicated at least three difficulties with this concept for our purposes: (1) neurotic behavior includes many different kinds of behavior and, hence, (2) the instruments used to measure it often measure different things, and (3) where there is agreement upon the definition of neurosis, the personality picture becomes too clinical, too extreme. The disgruntled, the disaffected recruits for radicalism described by Mayo, Lasswell, Santayana, *et al.* need not be the obsessive-compulsives of the clinic.

What, perhaps, is common to the hypothetically defined, socially gauche convert to radicalism, as well as to the more seriously ill neurotic, is anxiety—a “tenseness-in-the-bowels” occasioned by people. There is discomfort with others, a feeling that social relations are crises, that people are

more often a nuisance than a comfort, that they are objects one has to contend with rather than enjoy.

When the testing literature was scanned for some measure of this quality, it seemed that the closest approximation was inherent in Maslow's conception of security-insecurity (26). The Maslow Security-Insecurity Inventory has been developed on college groups and, despite its rather obvious and unprojective nature which makes it susceptible of faking, it has a satisfactory reliability and validity (27). It has been validated against interview ratings, self-ratings, and proclivity to seek psychotherapeutic counseling. All these validating checks were made on college samples and are in this wise limited. A 25-item subscale of Maslow's 75-item test was used, scored from 0 (secure) to 25 (insecure).

The R-C Scale

A measure of political-economic opinion was constructed by the usual item-winning procedure (49) with some items borrowed with permission from Eysenck (10, 11), Centers (7), and the California Opinion Study (1). The final scale of 14 items² had a split-half, corrected reliability on a two-

² The final items are:

1. Ultimately, private property in the instruments of production should be abolished and complete socialism introduced.

2. Profits of the great industries should be rigidly controlled by the Federal government.

3. The nationalization of the great industries is likely to lead to inefficiency, bureaucracy, and stagnation.

4. The reason that many advocate "free enterprise" is because it will enable them to continue exploiting the workers.

5. The traditional capitalistic system provides for the best possible distribution of wealth, human nature being what it is.

6. Labor unions should become stronger and have more influence generally.

7. The right to inherit wealth is a sound principle which provides a strong incentive for creative work.

8. In a socialist system the worker maintains his dignity and self-respect, while under capitalism he is just a tool or instrument to be exploited.

9. In general, full economic security is harmful; most men wouldn't work if they didn't need the money for eating and living.

10. No one should be allowed to earn more than \$25,000 a year "take-home" income.

11. It is up to the government to make sure that everyone has a secure job and a good standard of living.

12. America may not be perfect, but the American way has brought us about as close as human beings can get to a perfect society.

13. Wages and salaries would be fairer, jobs more steady, and we would have fewer people out of work, if the government took over and ran our mines, factories, and industries.

14. The present arrangement for the distribution of wealth is altogether unsound.

TABLE 1
Political Opinion Scores of "Known Groups"

Political Group	Mean Score	SD
Republicans (44)	61.8	5.4
Democrats (14)	36.3	4.0
Socialists (15)	6.4	4.6

week retest of $+.876 \pm .09$ on a sample of 113 upperclassmen and graduate students at the University of California, Berkeley, 1950. In administration, a six-point, Likert-type (24) forcing rating is used, scored 0-5, and making the range of scores for the 14 items 0-70, radical to conservative.

Efforts to validate the scale on "known groups" (29) met with the usual resistance on the part of many nonacademic subjects. Nevertheless, it was possible to obtain the cooperation of sufficient "known groups" to assure the validity of the political-economic scale. Responses were obtained from 44 members of the Republican Central Committee of Los Angeles County, from 14 active workers in the Democratic party headquarters in Santa Barbara, and from 15 members of the Socialist party in California.³

Table 1 indicates the wide range of mean political-opinion scores obtained from these groups. The differences between means are obviously highly significant and in the expected direction. (Unless otherwise stated hereafter, differences between means and departures of distributions from chance patternings are called significant only at the .01 or better confidence level.)

A secondary check on the validity of our measure of radicalism-conservatism is gained by comparing the political party affiliations of our entire sample of 538 respondents (cf. Table 2) with scores made on the attitude scale. The distribution of scores by party connection differs greatly from chance expectancy with Republicans scoring conservatively, Socialists and Progressives radically, and Democrats between them. The distribution also conforms with the expected association of indifference and neutrality remarked by Guttman and others (12, 15, 16). The majority of "no responses," "decline-to-states," and "miscellaneous party" affiliates falls in the middle, less intense range of scores.

³ We wish to express our appreciation of the ready assistance given us by Mr. Orville J. Evans, executive secretary of the Republican Central Committee of Los Angeles and Mr. Frank Jorgensen, vice-chairman of that committee. Mrs. John Jennings was most helpful in obtaining responses from workers in the Santa Barbara County Democratic Central Committee of which she was vice-chairman. We wish also to thank Miss Lillian Muniz, acting secretary of the Socialist Party, local Los Angeles, for distributing copies of our questionnaire.

One response from the National Council of the Arts, Sciences, and Professions (NCASP) is not included among the validating scores, but will be found in the total tabulations, Table 2.

A third and minor validation of our scale was obtained by comparing radical-conservative scores with self-ratings. Subjects rated themselves on a left-right continuum as very radical, progressive, radical, mildly radical, liberal, mildly conservative, conservative, or very conservative. Despite the obvious possibility of semantic confusion and inversion of the suggested steps on the self-rating continuum, there is a significant association between attitude score and self-rating.⁴

The Scale Referent

This investigation was begun with the assumption that political and economic attitudes are related through their association with the central concept of *power*, this being defined simply as the ability to get what one wants in a communal situation (51). And, while there are many aspects of society toward which disgruntled attention may be turned, prevailing ideologies focus on the power complex, and our items refer to different facets of this complex: "Who should get how much of whatever there is to get, and how?" This is the principal referent of our scale within the context of what G. H. Smith (42) has called "the central issue of all today's politico-economic strife: two opposed philosophies of government, an individualistic one and a collectivistic one."

The validation of our political opinion scale assures us that we are measuring along this continuum with sufficient discrimination for purposes of group comparison. Further tests of unidimensionality or homogeneity were, therefore, not felt to be required.⁵

THE SUBJECTS

In order to assure a greater range of responses, attempts were made to gather data from nonacademic subjects. In addition to the validating groups mentioned above, responses were obtained from 13 executives of the International Association of Machinists, 7 officers of the International Hod-

⁴ Our findings do not agree with those of Rice (35) who found a proclivity of college students in 1926 to rate themselves "conservative-liberal" or "liberal-conservative." Some 82 per cent of his sample chose such self-ratings. Our subjects are far less reluctant to label themselves radical or conservative.

⁵ The as yet unsettled status of different measures of unidimensionality—factor analysis, scale analysis, and the technique of homogeneous tests—also justifies our reluctance to enter the "pure test" fray.

On this score, however, we cannot agree with Guttman's parody of known-group validation (45). It is not true, as he alleges, that known group validation involves selecting groups without defining the attitude in advance. Nor is it true, as he states, that the kinds of behavior to be expected from these attitudes is not known, or cannot be made explicit, by the known group approach. For political opinions, voting behavior provides such a validating check; for ethnic opinions, the *action* of such bodies as the Hood River Legionnaires, Oregon, 1942-1945, is an example of such validation (29).

TABLE 2

Means and Standard Deviations of Political Opinion (R-C) and Security Scores (S-I) by Sample

Sample	N	R-C		S-I	
		\bar{x}	SD	\bar{x}	SD
Known Groups					
Republicans	44	61.8	5.4	4.2	3.6
Democrats	14	36.3	4.0	5.9	3.0
Socialists	15	6.4	4.6	5.3	3.2
Nonacademic Subjects					
Miscellaneous	64	49.5	13.1	5.4	4.7
Machinists	13	38.2	11.3	5.2	2.7
Hodcarriers	7	46.6	7.2	7.1	1.1
ILGWU	10	21.2	11.4	7.7	4.4
NCASP	1	12.0	—	13.0	—
College Students					
USC	76	44.9	12.5	5.1	3.7
Reed College	54	29.0	10.5	10.0	5.3
UC (Santa Barbara)	29	40.4	8.1	6.7	4.6
Stanford Univ.	170	48.4	10.1	5.7	4.3
Santa Barbara PN's	21	42.3	10.4	7.9	4.6
Santa Barbara Control	20	40.5	8.6	5.4	4.5
Total	538	37.0	9.0	6.8	3.8

carriers' and Building Laborers' Union, and 10 executives of the International Ladies' Garment Workers' Union—all Los Angeles locals.⁶ Sixty-four "miscellaneous" adults living in Los Angeles and Santa Barbara also completed questionnaires.

The bulk of our sample, however, consists of college students: 76 sociology students at the University of Southern California, 54 students from Reed College, 29 from the University of California at Santa Barbara, and 170 students in the Graduate School of Business at Stanford University.⁷ A group of 21 Santa Barbara College students rejected for teacher training as psychoneurotic because of excessively high scores on the "neurotic triad" of the MMPI was compared with a roughly matched group of 20 controls from the same college. Table 2 lists the means and standard deviations of scores obtained by these groups on our measures of radicalism-conservatism and security-insecurity.

⁶ We are grateful to Mr. Ray Waters of the Hodcarriers' Union, Mr. Edward Skagen of the Machinists' Union, and Mr. Sigmund Arywitz of the ILGWU for their cooperation.

⁷ Professors J. M. Willits of Stanford University, Melvin Vincent of the University of Southern California, and Howard Jolly of Reed College were most helpful in distributing our opinionnaires.

RESULTS

Security Feelings and Political Opinion

A significant association appears between security and conservatism, insecurity and radicalism, as measured. The Pearsonian coefficient between insecurity and conservatism is $-.51 \pm .043$. The relationship, however, is curvilinear and higher than that indicated by this rectilinear measure. The correlation ratio, *etc.*, computed with security scores as the independent variable and radicalism-conservatism as the dependent function, is .618. This is a highly significant departure from rectilinearity as measured by the χ^2 test.

Inspection of the scatterplot of radical-conservative (R-C) and security-insecurity (S-I) scores tells us that our sample suffers from a lack of extremely radical and extremely insecure subjects. The sample is predominantly secure and conservative—secure and conservative, that is, in terms of the possible range of scores, not in terms of any established norms for there are, as yet, none such.

At first blush, the rather restricted range of obtained scores should encourage us to assume that widening the range—adding more insecure and more radical subjects—would increase the degree of association. However, we are reluctant to make this assumption for the following reasons:

1. The principal contribution to the association of R-C and S-I scores comes from the large number of secure conservatives, rather than from an association of insecurity with radicalism.

2. The association which is found between insecurity and radicalism varies with the type of radical. Radicalism seems related to insecurity only when it is the radicalism of the "lone wolf." Our data indicate that the individuated, "degroupped" radical is apt to score insecurely. The Reed College sample provides a paradigm of the unorganized and insecure dissenter (Cf. Table 2).

However, the group-protected radical, the ideologue of the left, scores as securely as his counterpart of the right. Our active Socialist party members are as secure as our Republicans. This would substantiate Diamond's genetic study (9) in which he argued that insecurity may lead to the acceptance of radical ideology and that membership in an active, leftist group seemed to provide a purposeful self-conception which relieved these feelings of *gaucherie*. Normative radicalism, such as is sometimes found in limited, isolated communities, need not be related to personal insecurity (31, 32).

In summary, the association between personal security and political opinion which obtains in our limited sample causes us to conclude that political conventionality, conformity, conservatism and expressed feelings of personal security tend to go together—particularly among those of informal political persuasion. When one examines this relationship among the politically active, the "group-grooved" and ideologically convinced, the

association between type of political belief and security vanishes. Instead, we find the politically active of all persuasions scoring in a very self-assured manner. The late anarchist philosopher, Simone Weil, would assent, "Armored by the lie" (53).

Substantiation of the relationship found between the direct measures of political opinion and personal security may be gained by noting how these variables are related indirectly—particularly through their association with such other correlates of political opinion as political party affiliation and social-class identification.

The χ^2 test applied to a table showing the distribution of security scores by political party membership or preference indicates the distribution to be nonchance, and examination of the table shows the discrepancies to be in the expected direction. A disproportionate number of Republicans score securely and less than their expected number score insecurely. The distribution of Democratic S-I scores is uneven, with a slight tendency for Democrats to receive less than their expected allotment of very secure responses and more than their share of very insecure scores. Socialists and Progressives have a slight excess of very insecure responses while people of miscellaneous persuasion and those who decline-to-state or indicate no-party membership score much less securely than expected and more insecurely than chance would allow.

The relationship between social-class identification and R-C and S-I scores is also nonchance and in the expected direction. Subjects who identify with the "professional and upper" class score conservatively, those who rate themselves as "upper middle" class obtain significantly less than the expected number of radical scores, and individuals who identify with the "lower or working" classes score radically.

And, as expected, our upper class affiliates score securely while those of lower or workingclass consciousness score insecurely.⁸

It is remarkable that, in obtaining class identifications in an open-ended manner, almost as many subjects are willing to signify class affiliation as political party membership. While some 86 respondents decline to identify with a social class, or say "don't know" or "don't believe in classes," or, as in four cases, give some aberrant identification such as "all classes" or "American," almost as many subjects, 77, refuse to identify with a political party.

⁸ Social class identification was obtained through an open-ended question, "In which social class, if any, do you consider yourself?"

We have long doubted the value of public opinion polls which have used multiple-choice items on this subject and which find most Americans identifying with the "middle class" (*Fortune* poll, 1940) or, as in Centers' revision (7), with the "working class." Although our sample is far from representative of the adult population, our results suggest that public opinion pollsters might come to different conclusions concerning social class identification if they used an open-ended question.

However, while those who refuse to affiliate with a political party tend to score radically, those who decline social class identification are more likely to be *either* radical or conservative. As expected because of the association of R-C and S-I scores, those who do not signify political party membership score insecurely. But those who decline class identification score less insecurely than would be expected by chance.

Additional corroboration of the correlation between political opinion and security revealed by the direct measures and the above indirect measures is gained from the relationship of these variables to political self-ratings. It has already been noted that R-C scale scores are significantly associated with self-ratings. These self-ratings also bear a relationship in the expected manner to security scores. However, this relationship is attenuated—possibly by the connotations involved in the rating labels—so that a similar association is to be expected by chance about three times in a hundred whereas the distribution of self-ratings and R-C scores deviates farther from chance expectation.

Nevertheless, our unpublished tables, available upon request, show the anticipated tendency of self-rated radicals to score less securely and more insecurely while self-labeled conservatives score more securely than by chance.

In summary, the hypothesis that conservatism is, in part, a function of security and radicalism of insecurity is substantiated by the correlation between our measures of these variables and by their association with political party membership, social class identification, and self-ratings of political position.

This is not to conclude that such social "dis-ease" as may be gauged by the Maslow S-I scale is an uncomplicated determinant of political opinion and security. It is entirely conceivable that holding a deviant political opinion may produce insecurity as well as reflect it, although we believe the principal flow is in the other direction.

The Status Variables

This reasoning gains justification from the supportive role of status variables—income, occupation, education, religion, and age. Since these variables are related to the R-C scale referent, power, their association with our R-C and S-I scores becomes important. The following summary is derived from tables unpublished for reasons of economy but available upon request. The χ^2 test is called significant where p 's = $<.01$; lesser deviations from chance distribution are indicated.

Income. The relationship of income to political opinion is highly significant whereas that between income and security score is of only "borderline" significance. (At about the .07 confidence level.)

It is not unusual to find, as we do, that the wealthy are conservative

and the poor, radical, but an additional, important contribution to the association between political opinion and income is made by a middle-income (\$3,000-\$5,000 per year)⁹ group of radicals. This finding is in keeping with the oft-noted recruitment of the radical intelligentsia from among the educated middle class, and would be anticipated in our highly educated sample.

The lesser relationship between income and security gains significance from the following associations:

- a. There are fewer secure poor individuals than would be expected if the distribution were random.
- b. There are more insecure poor than expected.
- c. There are more secure well-to-do (\$5,000-\$7,500) than expected.
- d. There are fewer insecure middle-income and well-to-do (\$3,000-\$7,500) than expected.

Occupation. As hypothesized, occupation is related very much in the same way as is income to political opinion and security, although the constricted range of occupations in our sample necessarily limits the association.

Nevertheless, there is a highly significant association between political viewpoint and occupation with the higher status occupations being the more conservative. The distribution of political opinions by occupation deviates from a merely random distribution largely because of the greater-than-chance number of:

- a. Radical clerks, salesmen, and kindred workers
- b. Radical skilled and semiskilled workers
- c. Conservative business executives and proprietors
- d. Conservative housewives
- e. The fewer-than-expected number of middle-of-the-road (R-C scores between 30-49) business executives and proprietors

The relationship of occupation to security score is of only marginal significance. A distribution such as we have obtained would be expected by chance about eight or nine times in a hundred and can be summarized as indicating that:

- a. Students score less securely and more insecurely than by chance expectancy.
- b. Professional workers and business executives score more securely and less insecurely than by chance.
- c. There are fewer secure clerks and salesmen than expected.

⁹ Median family income in the United States has been reported for 1949 as \$3,100. For that year approximately 18 per cent of American families had income of over \$5,000. per year and about 8 per cent received more than \$7,500., according to a Bureau of the Census survey, March, 1950.

- d. There are fewer insecure skilled and semiskilled workmen than by chance allowance.

Education. With some exceptions, the conclusion of many American studies on the role of education in the formation of political attitudes favors the "liberalizing" influence of formal schooling.¹⁰ However, our results do not uniformly support this traditional view. Our college graduates and those in postgraduate study score more conservatively than by chance expectancy. But some of this tendency is determined by the conservative sample from the Stanford Graduate School of Business and these results certainly cannot be generalized. Our subjects with *some* college training score more radically than by chance.

The association we find between high education and conservatism is coupled with an association of high level education and security. Score distributions indicate the lesser security of college students, the greater security of our college graduates and postgraduate students, and the moderate security level of those with twelve years or less education.

Religion. A clearer relationship between a status-symbol and conservatism-security is revealed by the association of these variables with religion. These associations may be summarized as follows:

a. *As regards radicalism-conservatism:*

- i. Protestants score conservatively and nonradically.
- ii. Catholics are conservative, but *not* nonradical.¹¹

¹⁰ Thus Jones (20). However, more recent studies cast doubt on the necessarily "liberalizing" influence of formal education. Cahalan and Trager (6) report that the relationship of education and socioeconomic status to anti-Semitism may depend upon the way in which attitudes toward Jews are measured. The typical attitude scale gives a negative relationship between education and prejudice, but "free-answer stereotyping" reveals a positive relationship. Perhaps the educated learn which attitudes are "nice," and how to respond on the ordinary questionnaire accordingly.

Similarly Smith (40, 41, 43) has shown the relationship of "liberalism" and education-information to depend upon the opinion area. When the issues are noneconomic and interclass in nature (e.g., UN vs. U.S. control of the atomic bomb), "liberalism" and superior general information and education are *positively* correlated. But when the issues are politico-economic in implication (e.g., government guarantees of jobs and standard of living vs. helping people get ahead on their own), "liberalism" and education-information are *negatively* correlated.

¹¹ Without the χ^2 tables before him, the statistically unsophisticated reader may be confused by such a seeming redundancy as "conservative and nonradical" and such an apparent contradiction as "conservative, but *not* nonradical."

For each cell of a χ^2 table, one can note whether subsamples contribute their randomly expected number of cases, or more or less. The contributions of subsamples toward deviation from a chance distribution of their numbers among the various categories need not be uniform. Thus "Protestants score conservatively" means that they had a greater number of conservative scores than expected from a random

- iii. Jews are radical and nonconservative (2, 46).
 - iv. Those of miscellaneous faith, or who decline to state, or who have no faith score as do the Jews.
 - v. The atheists and agnostics are radical, but not necessarily nonconservative.
- b. As regards security-insecurity:*
- i. Protestants are less insecure than by chance.
 - ii. Catholics are about as secure and insecure as chance alone would allow.
 - iii. Jews show a slight tendency toward insecurity.
 - iv. Atheists and agnostics score similarly to Jews.
 - v. The greatest contribution to the nonchance association of religion and security stems from the lesser security and greater insecurity of those of miscellaneous faith or no faith.

Age. Although age is not so strong a mark of status in our society as are income, occupation, education, and religion, we thought it of interest to examine the association of age with expressions of security and political opinion. The relationships are very low but statistically significant and in the expected direction. Age and security correlate $+ .14 \pm .014$ and age and conservatism $+ .28 \pm .014$.

Our findings, then, lend support to the notion that status is associated with conservatism and expressed feelings of security and such a finding is entirely in keeping with the classic hypothesis of the resentful radical.

Intensity and Neurotic Behavior

It was indicated earlier that our first intention was to relate political opinion to neuroticism, but that we concluded that the concept of neurosis was either too vague or, where clarified, too clinical to describe the type of feeling our literary hypothesis had pictured as motivating radicalism.

Nevertheless, we were interested in investigating the relationship between neurotic behavior, insecurity, and political opinion. A sample of all students at Santa Barbara College rejected as teacher candidates because of excessively high scores on the "neurotic triad" of the Minnesota Multiphasic Personality Inventory was compared with a control group of non-neurotic students. The political opinion and security scores of these two groups are given in Table 2.

distribution. This says nothing about their contribution to radical scores; the distribution of R-C scores need not be rectilinear. Hence it is meaningful to add that Protestants score nonradically, i.e., that they contribute *less* than their expected number to the radical cell.

Contrast this pattern with the Catholics who also score conservatively, but who place their expected number of subjects in the radical cell. Hence, they are "*not* nonradical."

There is no statistically significant difference between either the R-C or S-I scores of these psychoneurotics and normals. Apparently the hysteria, depression, and hypochondriasis triad of the MMPI is tapping a different behavioral area than Maslow's S-I scale. (For a larger sample of high-school seniors, Gough (13) found a moderate correlation between the full MMPI and S-I scales. The correlations, all positive, were .186 with H_1 , .400 with D , and .224 with H_2 .)

However, there is a tendency of our psychoneurotics to gain their political opinion scores by making more extreme responses. They seem to hold their views with greater intensity; they are more willing to respond to statements with "strongly agree" or "strongly disagree."

But while there is this apparent tendency for our Santa Barbara neurotics to score more intensely, when we examine the relationship between security scores and intensity for our entire sample, the relationship is in an opposed direction; there is a low, but significant, association between intensity and security ($+ .14 \pm .014$).

Sex, Security, and Political Opinion

Studies of political opinion usually note the greater conservatism of women. This finding does not agree, within the context of the present hypothesis, with the usual finding of greater neuroticism among women, nor with our results.

The average R-C score of the women in our sample is reliably more radical than the average male score, and this difference yields a point biserial coefficient between sex and political opinion of .22. However, there is no significant difference between the mean security scores of the sexes.

These findings are, of course, again qualified by the unrepresentative nature of our sample.

CONCLUSION

Are those persons who feel insecure more likely to look with favor upon a change of the power-wealth complex in the United States?

The answer, within the limits of our instruments and sample, is affirmative. When 538 academic and nonacademic subjects were tested with reliable and valid measures of political-economic opinion and security, a curvilinear association was found ($Eta = .618$).

This relationship means that secure people tend to have conservative political opinions. The principal exception to this occurs in the case of those radicals who are participating members of active political organizations. Such "group-armored" radicals score very securely. The insecure radical may be characterized as the man of little faith—political or religious; he is degroupped, individuated, *deraciné*. He is the marginal man who dis-

likes the existing order but who has not yet gained the security that flows from participating with others of like mind in a "cause."

While the association of insecurity and radicalism cannot in itself be interpreted causally and while the possibility of interaction between these attitudinal complexes is acknowledged, still this study lends support to the traditional hypothesis that the motivational sequence is from personal to societal attitudes, rather than vice versa.

The central assumption that those who feel insecure personally are more susceptible of radicalism is also supported by the relationship between certain status indices and political opinion-security scores. High status as represented by income, occupation, education, religion, and, to a lesser extent, age, is associated with conservatism and security; low status with radicalism and insecurity.

Manuscript received: November 1, 1955

Revised manuscript received: November 5, 1956

Gwynne Nettler
Monterey Peninsula College
Monterey, California

REFERENCES

1. Adorno, T. W., et al., *The Authoritarian Personality*, New York: Harper, 1950.
2. Allin Smith, W., and B. Allin Smith, "Religious Affiliation and Politico-economic Attitudes: A Study of Eight Major U.S. Religious Groups," *Public Opinion Quarterly*, 1948, 12, 377-389.
3. Allport, F. H., and D. A. Hartman, "Measurement and Motivation of Atypical Opinion in a Certain Group," *American Political Science Review*, 1925, 19, 735-760.
4. Anastasi, A., and J. P. Foley, *Differential Psychology*, New York: Macmillan, 1949, Chap. 23.
5. Bray, D. W., "The Prediction of Behavior from Two Attitude Scales," *Journal of Abnormal and Social Psychology*, 1950, 45, 64-84.
6. Cahalan, D., and F. N. Trager, "Free Answer Stereotypes and Anti-Semitism," *Public Opinion Quarterly*, 1949, 13, 93-104.
7. Centers, R., *The Psychology of Social Classes*, Princeton, N. J.: Princeton University Press, 1949, particularly Chap. 5.
8. Dexter, E. S., "Personality Traits Related to Conservatism and Radicalism," *Character and Personality*, 1939, 7, 230-237.
9. Diamond, S., "A Study of the Influence of Political Radicalism on Personality Development," *Archives of Psychology*, 1936, 203.
10. Eysenck, H. J., "General Social Attitudes," *Journal of Social Psychology*, 1944, 19, 207-227.
11. Eysenck, H. J., "Primary Social Attitudes, I. The Organization and Measurement of Social Attitude," *International Journal of Opinion and Attitude Research*, 1947, 1, 49-84.
12. Foa, U. G., "Scale and Intensity Analysis in Opinion Research," *International Journal of Opinion and Attitude Research*, 1950, 4, 192-208.

13. Gough, H. G., "A Note on the Security-Insecurity Test," *Journal of Social Psychology*, 1948, 28, 257-261.
14. Gundlach, R., "Emotional Stability and Political Opinions as Related to Age and Income," *Journal of Social Psychology*, 1939, 10, 577-590.
15. Guttman, L., "Suggestions for Further Research in Scale and Intensity Analysis of Attitudes and Opinions," *International Journal of Opinion and Attitude Research*, 1947, 1, 30-35.
16. Guttman, L., and E. A. Suchman, "Intensity and a Zero Point for Attitude Analysis," *American Sociological Review*, 1947, 12, 57-67.
17. Howells, T. H., "A Comparative Study of Those Who Accept as Against Those Who Reject Religious Authority," *University of Iowa Studies*, 1928, 2, 2.
18. Janet, P., *Les Neuroses*, Paris: E. Flammarion, 1915.
19. Jones, E., "Evolution and Revolution," *International Journal of Psychoanalysis*, 1941, 22, 193-208.
20. Jones, V., "Attitudes of College Students and the Changes in Such Attitudes during Four Years in College," *Journal of Educational Psychology*, 1938, 29, 14-25 and 114-134.
21. Jung, C. G., *Psychological Types*, New York: Harcourt, Brace, 1923.
22. Krout, M. H., and R. Stagner, "Personality Development in Radicals," *Sociometry*, 1939, 2, 31-46.
23. Lasswell, H. D., *Psychopathology and Politics*, Chicago: University of Chicago Press, 1930.
24. Likert, R., "A Technique for the Measurement of Attitudes," *Archives of Psychology*, 1942, 140.
25. Lynd, R. S., "The Motivation of the Left-Winger," in A. M. Lee and E. B. Lee (eds.), *Social Problems in America*, New York: Holt, 1949.
26. Maslow, A. H., "The Dynamics of Psychological Security-Insecurity," *Character and Personality*, 1942, 10, 331-344.
27. Maslow, A. H., et al., "A Clinically Derived Test for Measuring Psychological Security-Insecurity," *Journal of General Psychology*, 1945, 33, 21-41.
28. Mayo, Elton, *The Social Problems of an Industrial Civilization*, Andover, Mass.: Andover Press, 1945.
29. Nettler, G., and E. H. Golding, "The Measurement of Attitudes toward the Japanese in America," *American Journal of Sociology*, 1946, 52, 31-39.
30. Newcomb, T. M., in G. Murphy, L. B. Murphy, and T. M. Newcomb, *Experimental Social Psychology*, New York: Harpers, 1937, p. 941.
31. Newcomb, T. M., *Personality and Social Change*, New York: Dryden, 1943.
32. Newcomb, T. M., "The Influence of Attitude Climate upon Some Determinants of Information," *Journal of Abnormal and Social Psychology*, 1946, 41, 291-302.
33. Nietzsche, F., *The Genealogy of Morals*, New York: Tudor, 1931.
34. Plant, J. S., *Personality and the Cultural Pattern*, New York: The Commonwealth Fund, 1937.
35. Rice, S. A., "Differential Changes of Political Preference under Campaign Stimulation" *Journal of Abnormal and Social Psychology*, 1926, 21, 297-303.
36. Sanai, M., and P. M. Pickard, "The Relation between Politico-economic Radicalism and Certain Traits of Personality," *Journal of Social Psychology*, 1949, 30, 217-227.
37. Santayana, G., *Dominations and Powers*, New York: Scribners, 1951.
38. Scheler, M., *L'Homme du Ressentiment*, Paris: Librairie Gallimard, 1933.

39. Schumpeter, J. A., *Capitalism, Socialism, and Democracy*, New York: Harper, 1947.
40. Smith, G. H., "The Relation of 'Enlightenment' to Liberal-Conservative Opinions," *Journal of Social Psychology*, 1948, 28, 3-17.
41. Smith, G. H., "Liberalism and Level of Information," *Journal of Educational Psychology*, 1948, 39, 65-81.
42. Smith, G. H., "Rejoinder to Hugh J. Parry," *Public Opinion Quarterly*, 1948-1949, 12, 784-786.
43. Smith, G. H., "Opinions Related to College Expectations and Vocabulary Level," *Journal of Social Psychology*, 1950, 32, 255-263.
44. Stagner, R., "Studies of Aggressive Social Attitudes. III. The Role of Personal and Family Scores," *Journal of Social Psychology*, 1944, 20, 129-140.
45. Stouffer, S. A., et al., *Studies in Social Psychology in World War II: Volume IV, Measurement and Prediction*, Princeton, N. J.: Princeton University Press, 1950, pp. 54-55.
46. Sward, K., and M. B. Friedman, "Jewish Temperament," *Journal of Applied Psychology*, 1935, 19, 70-84.
47. Terman, L. M., *Psychological Factors in Marital Happiness*, New York: McGraw-Hill, 1938, pp. 145-146.
48. Thomas, N., *A Socialist's Faith*, New York: Norton, 1951.
49. Thurstone, L. L., and E. J. Chave, *The Measurement of Attitude*, Chicago: University of Chicago Press, 1929.
50. Toynbee, A., *A Study of History*, New York: Oxford, 1947, one volume edition, p. 377.
51. Weber, M., "Class, Status, Party," in H. H. Gerth and C. W. Mills, *From Max Weber: Essays in Sociology*, New York: Oxford, 1946, p. 180.
52. Weber, M., "Politik als Beruf," *Gesammelte Politische Schriften*, Munich, 1921, translated by H. H. Gerth and C. W. Mills in *From Max Weber: Essays in Sociology*, New York: Oxford, 1946, p. 125.
53. Weil, S., "The Iliad or The Poem of Force," *Politics*, November, 1945, 2, 321-331.
54. Wolfe, A. B., *Conservatism, Radicalism, and Scientific Method*, New York: Macmillan, 1923.

Sibling Structure and Perception of the Disciplinary Roles of Parents¹

ANDREW F. HENRY, *Vanderbilt University*

This paper reports findings supporting the hypothesis that perception by children of the disciplinary role structure within the family varies with their order of birth.

Recent research has shown that cardiovascular reactions of male subjects during experimentally induced stress vary with the subject's perception of the disciplinary role of the mother as compared with the disciplinary role of the father. Subjects responding with a cardiovascular reaction similar to that produced by infusion of epinephrine perceive their mother as playing the principal disciplinary role. Those responding with a cardiovascular reaction similar to that produced by infusion of norepinephrine perceive the father in the principal disciplinary role (4).

Type of cardiovascular reaction during experimentally induced stress also is related to the direction of expression of anger. Subjects reacting to experimental frustration with an outward display of anger against the experimenter experience a norepinephrinelike cardiovascular reaction. Those who blame themselves for their frustrating experience, or react with anxiety, experience an epinephrinelike cardiovascular reaction (3). Since epinephrinelike cardiovascular reaction was found to be associated (a) with the turning of anger aroused by frustration inwardly against the self and (b) with perception of mother as principal disciplinarian, the suggested hypothesis that intropunitive *attitudes* would be associated with perception of the mother as principal disciplinarian has been recently tested.

In a sample of 765 enlisted males in the U. S. Air Force, subjects saying they would blame themselves in hypothetical interaction situations involving harm to another person were more likely to perceive the mother in the principal disciplinary role than subjects saying they would not blame themselves (2).

These studies provide considerable evidence of the existence of a relation between perception of mother as principal disciplinarian and intropuni-

¹ Revision of a paper read before the annual meeting of the American Sociological Society at Detroit, September, 8, 1956. The writer wishes to acknowledge with thanks the cooperation of each of the following in facilitating the collection and processing of data: Superintendent of Schools, Fall River, Massachusetts; Superintendent of Schools, Malden, Massachusetts; Principal of Goodlettsville High School, Goodlettsville, Tennessee; Principal of DuPont High School, Old Hickory, Tennessee; Joseph T. Drake, Department of Sociology, University of Tennessee; David Bordua, University of Michigan; and James Williams, Louis Dotson, Edward MacDill, Jean Dedman, Eunice Ribeiro, and Dorothy Griffin, Vanderbilt University.

tiveness measured on both the affective-behavioral and cognitive-attitudinal levels. When the mother is perceived as the principal disciplinarian, anger consequent to frustration is inhibited and turned inwardly against the self. When the father is viewed as disciplinarian, anger flows outwardly against others.

The Gluecks present substantial evidence that delinquent males discharge their anger outwardly against others more than nondelinquents with whom they are matched. From their psychiatric interviews, they conclude that 67.5 per cent of delinquents as compared with only 30.5 per cent of nondelinquents resolve conflict by "extroversion" of affect or action. "Extroversion of action involves a direct, behavioristic resolution of tension; . . . Extroversion of affect implies . . . an indulgence of feelings and instincts . . ." (1, p. 251). From their Rorschach data, they find that delinquents are more defiant, have more feelings of resentment, evidence greater hostility, and manifest fewer "masochistic trends" than nondelinquents (1, pp. 219, 225, 228, 234).

Since males who discharge their anger outwardly perceive father as principal disciplinarian more than those who inhibit their anger, we would expect delinquents to perceive father as disciplinarian more than nondelinquents. Data from the Gluecks' study indicate that this is the case. They rate 56.8 per cent of mothers of delinquents as lax in discipline as compared with 26.6 per cent of fathers. Among the delinquents, the father plays the principal disciplinary role. Among the nondelinquents the reverse is true. Only 11.7 per cent of the mothers of nondelinquents are rated lax in discipline as compared with 17.9 per cent of the fathers (1, p. 131).

These data all combine to suggest that perception of father as principal disciplinarian is associated with the outward discharge of anger. The relation holds when we index the direction of flow of aggression psychophysiological, when we index it attitudinally, and when we index it by comparing delinquents with nondelinquents. Those reacting during experimental stress with an outward discharge of anger perceive father as disciplinarian more than those turning experimentally induced anger upon themselves. Those saying they would blame others in hypothetical situations posed on attitude surveys perceive father as disciplinarian more than those saying they would blame themselves. Delinquents who tend to be outwardly aggressive come disproportionately from families where the father is the principal disciplinarian and the mother is lax in discipline.

Sletto, back in 1934, demonstrated that delinquent children come disproportionately from families in which there are younger siblings. "In general, delinquency ratios were high for children who are in sibling positions involving the presence of younger siblings of each sex, and low for children in positions involving the presence of elder siblings of each sex" (5, p. 665). Since delinquents, like others expressing their anger outwardly, perceive

father as principal disciplinarian, Sletto's finding suggests the hypothesis that perception of father as disciplinarian, like delinquency, should be concentrated among the first born. If it should turn out that the eldest child perceives father as disciplinarian while the youngest child perceives mother as disciplinarian, we may suggest that the relation of delinquency to birth order found by Sletto could be accounted for by the relation of the disciplinary structure to birth order.

A relation between disciplinary structure and order of birth is somewhat easier to rationalize than a relation between delinquency and order of birth. In infancy, the mother in our culture assumes responsibility for the care of the child. As the child ages, the father shares increasingly in his nurture and discipline. Our problem, however, is not concerned with the distribution of discipline between mother and father of the first child at any particular age, but rather with the change in the distribution of discipline brought about by the arrival of additional children. The mother with a new baby is again forced to concentrate her attention on the infant. She has less time for nurture and discipline of the older child. This may operate to shift the burden of discipline of the older child over to the father. If this transfer of the disciplinary function to the father occurs systematically with increase in family size, we should find the oldest child in the family perceiving father as principal disciplinarian and the youngest child in the family perceiving mother as principal disciplinarian.

RESEARCH DESIGN

The hypothesis was tested on two samples. The New England sample, consisting of 1335 students in two Massachusetts public high schools ranged in age from thirteen to twenty-one with about 75 per cent falling within the fifteen to seventeen age group. The sample is heavily Catholic with 63 per cent falling in this religious category. Eighteen per cent of the subjects reported their fathers had either attended or graduated from college.

The Southern sample, composed of 391 college students and 226 high-school students in Tennessee ranged in age from fifteen to twenty-five with 75 per cent falling within the sixteen to twenty age group. The Southern sample is heavily Protestant with 86 per cent falling in this religious group. Forty-one per cent of subjects in the Southern sample reported their fathers had either attended or graduated from college. Subjects in both samples came from families in which both parents were living together.

RESULTS

The New England Sample

Each respondent was asked this question: "Who is the principal disciplinarian in your family?" Three answer categories were provided: mother, father, or other. Five per cent either left the question blank or checked

TABLE 1

Percent Perceiving Mother as Principal Disciplinarian by Order of Birth Holding Sex and Grade in School Constant: Unbroken Homes with at Least Two Children

Birth Order	Males		Females	
	9th & 10th Grades	11th & 12th Grades	9th & 10th Grades	11th & 12th Grades
Oldest child	.28 (89)	.43 (89)	.41 (87)	.39 (88)
Middle child	.33 (79)	.35 (79)	.53 (95)	.46 (87)
Youngest child	.49 (67)	.52 (67)	.47 (79)	.58 (89)

TABLE 2

Percent Perceiving Mother as Principal Disciplinarian by Order of Birth Holding Sex, Grade in School, and Sibling Sex Composition Constant: Unbroken Homes with at Least Two Children

Grade in School and Birth Order	Males			Females		
	Same Sex Siblings Only	Cross Sex Siblings Only	Both Male and Female Siblings	Same Sex Siblings Only	Cross Sex Siblings Only	Both Male and Female Siblings
9th & 10th Grades						
Oldest child	.34 (29)	.12 (16)	.29 (44)	.41 (22)	.36 (33)	.47 (32)
Middle child	.23 (13)	.30 (10)	.36 (56)	.67 (6)	.61 (13)	.50 (76)
Youngest child	.45 (20)	.52 (23)	.50 (24)	.52 (23)	.62 (24)	.31 (32)
11th & 12th Grades						
Oldest child	.52 (25)	.40 (40)	.37 (24)	.29 (28)	.48 (25)	.40 (35)
Middle child	.43 (21)	.36 (11)	.32 (47)	.59 (17)	.29 (7)	.44 (63)
Youngest child	.58 (24)	.43 (21)	.54 (22)	.64 (28)	.50 (28)	.61 (33)

"other" and these cases were removed from the analysis. Subjects were asked whether they had an older brother, an older sister, a younger brother or a younger sister. Each of the four questions was answered with Yes or No. These answers were cross-tabulated and form the basis of the analysis which follows. Size of family was not controlled.

Table 1 presents the proportion perceiving mother as principal disciplinarian by order of birth holding sex and grade in school constant. In each of the four tests, the youngest child is more likely to perceive mother as disciplinarian than the oldest child. And in two of the four cases (males in the 9th and 10th grades and females in the 11th and 12th grades) the 2 by 2 χ^2 's are significant beyond the .01 level. An additional variable is controlled in Table 2 showing the proportion of each sex perceiving mother as principal disciplinarian by order of birth holding grade in school and the sibling sex composition constant. In 11 out of 12 possible tests, the youngest child is

TABLE 3

Percent Perceiving Mother as Principal Disciplinarian by Order of Birth Holding Sex of Subject and Family Size Constant: Unbroken Homes

Birth Order	One Male Sib	One Female Sib	Two Sibs	Three Sibs	Four or More Sibs	Total
Males						
Oldest child	.16 (32)	.43 (28)	.27 (33)	.07 (14)	.20 (10)	.25 (117)
Middle child			.22 (23)	.29 (24)	.45 (20)	.31 (67)
Youngest child	.37 (16)	.43 (14)	.39 (18)	.50 (6)	.25 (8)	.39 (62)
Females						
Oldest child	.47 (36)	.43 (21)	.47 (32)	.27 (15)	.56 (9)	.44 (113)
Middle child			.41 (22)	.27 (30)	.50 (18)	.37 (70)
Youngest child	.55 (20)	.59 (17)	.67 (18)	.56 (9)	.40 (10)	.57 (74)

more likely to perceive the mother in the principal disciplinary role than the oldest child. This result is significant by the two-sided sign test at the .01 level. For example, 12.5 per cent of eldest males with sisters but no brothers (cross sex siblings only) perceive mother as principal disciplinarian. Fifty-two per cent of youngest children in the same category perceive mother as disciplinarian.

The relation between perception of disciplinarian and birth order also was tested, holding constant sex and education of mother rather than sex and grade in school. The pattern does not change. In each of the four tests, the youngest child is more likely to perceive mother as disciplinarian than the oldest child and all four of the 2 by 2 χ^2 values are significant beyond the .20 level. When the sibling sex composition is brought under control along with sex of the subject and education of the mother, in 11 out of the 12 possible tests, the youngest child is more likely to see mother in the principal disciplinary role than father.

The New England data lend support to the hypothesis that the distribution of the disciplinary function between mother and father changes with increase in the size of family. But certain qualifications must be noted. First, since the sample is heavily Catholic, we are in all probability dealing with families larger than the average American family. Since size of family was not controlled, we cannot reject the hypothesis that our results are a spurious function of family size. Fortunately, we were able to test the relationship holding family size constant in the Southern sample.

The Southern Sample

Six per cent either failed to answer the question "Who is the principal disciplinarian in your family" or checked "other" rather than mother or

father. These, together with a few saying the mother was regularly employed outside the home, were removed from the analysis.

Table 3 presents the proportion perceiving mother as principal disciplinarian by order of birth holding sex and family size constant. In 10 of the 12 independent tests, the per cent perceiving mother as disciplinarian is higher for the last born than for the first born. In one instant, the per cent viewing mother as disciplinarian among the eldest and youngest is equal. The hypothesis is denied in only 1 out of the 12 tests. This result is again significant beyond the .01 level by the two-sided sign test.

DISCUSSION

Our hypothesis that the eldest child tends to perceive father as principal disciplinarian while the youngest child views mother as principal disciplinarian is confirmed in two separate tests. It is supported in a heavily Catholic sample of New England high-school students with fathers of relatively low education, and in a heavily Protestant sample of Southern college and high-school students with fathers of relatively high education. The fact that only a very small proportion of the variance in perception of the disciplinary structure is accounted for by birth order may be in part a function of the fact that we have been comparing youngest and oldest children from different families. Birth order should account for a very substantial part of the within-family variance when siblings in the same family are compared with each other.

In 12 of 15 tests on the New England sample (Table 2), females are more likely to perceive mother as disciplinarian than males. This relation also holds in 12 of 13 tests on the Southern sample (Table 3). This suggests that mothers tend to punish daughters while fathers tend to punish sons. It seems probable, therefore, that the sibling sex distribution would also account for part of the unexplained variance.

We have noted that perception of the distribution of disciplinary roles within the family is associated with the direction of expression of anger in its psychophysiological, attitudinal, and behavioral forms. Data reported here suggest that this perception is related to order of birth and reflects a shift in the discipline of the older child from mother to father with arrival of a new child. But this does not explain why discipline by the father leads to the outward discharge of anger and delinquency while discipline by the mother leads to intropunitiveness, inhibition, and self-control. It is tempting to speculate that the relation occurs because mothers and fathers differ in the type of discipline which they administer. The Gluecks, for example, find that parents of delinquents use more physical punishment and less "reasoning" than parents of nondelinquents. Yet they also find that the per cent of mothers using physical punishment in their nondelinquent sample is ex-

actly the same as the per cent of fathers using physical punishment (1, p. 132).

The amount of discipline, as well as its distribution between mother and father, may prove to be relevant for control of the outward expression of anger. A child freed from any external control over aggression may find it difficult to establish internal or self-control over his anger. Further, the eldest child may conceivably be subjected to less total discipline than the youngest. Yet, if this explanation were correct, we would be forced to take the position that children who perceive their father as principal disciplinarian have been subject to a lesser total amount of discipline than children who perceive mother as principal disciplinarian. Evidence to support this position is lacking.

There is evidence, however, to indicate that women are less outwardly aggressive in American society than men. This suggests that the mother may typically place more emphasis on the importance of controlling aggression in the child than the father. When she does the disciplining, the child may be punished for aggressive outburst. When father is the disciplinarian, aggressive behavior in the child may go unpunished or even be rewarded.

An alternative hypothesis lies in the degree to which punishment by the mother as compared with punishment by the father results in a threat to the flow of love and nurture to the child. In our culture, the mother is the principal source of nurture. When the source of frustration and the source of nurture are centered in the same parental role as in the case when mother is principal disciplinarian, the outward discharge of anger against the source of frustration may threaten the flow of nurture leading to inhibition and control of anger. When these functions are split between parental roles, the child can at the same time receive love from the nurturing mother and discharge anger against the frustrating father. Anger does not jeopardize the flow of nurture and therefore is not inhibited.

SUMMARY

Psychophysiological, attitudinal, and behavioral data indicate that the outward discharge of anger is associated with perception of father as principal disciplinarian, while intropunitiveness is associated with perception of mother as principal disciplinarian.

Sletto's research showing that delinquents come disproportionately from families in which there are younger siblings suggested the hypothesis that perception of father as disciplinarian, like delinquency, should be concentrated among the first born. This hypothesis was tested in two samples and supported in both. Findings suggest that the disciplinary function of the older child is transferred from mother to father as family size increases.

Certain hypotheses to account for the relation between the structure of disciplinary roles and the direction of expression of anger are advanced.

Manuscript received: August 14, 1956

Revised manuscript received: October 11, 1956

Andrew F. Henry
Department of Sociology and Anthropology
Vanderbilt University
Nashville 5, Tennessee

REFERENCES

1. Glueck, S. and E., Glueck, *Unraveling Juvenile Delinquency*, New York: The Commonwealth Fund, 1950.
2. Henry, A. F., "Family Role Structure and Self Blame," *Social Forces*, 1956, 35, 34-38.
3. King, S. H., "Emotional and Cardiovascular Responses during Stress—An Experimental Study," Ph.D. dissertation, Harvard University, 1953.
4. King, S. H., and A. F. Henry, "Aggression and Cardiovascular Reactions Related to Parental Control Over Behavior," *Journal of Abnormal and Social Psychology*, 1955, 50, 206-210.
5. Sletto, R. F., "Sibling Position and Juvenile Delinquency," *American Journal of Sociology*, 1934, 39, 657-669.

Changes in Social Perception as a Function of The Personal Relevance of Behavior¹

EDWARD E. JONES, *Duke University*

RICHARD DECHARMS, *University of North Carolina*

The present investigation is an outgrowth of an earlier stated conception of social perception (3) which stressed some conditions under which different inferences will be drawn from the same behavior of another. If the behavior of a stimulus person has clearly defined consequences for the value-maintenance or the goal attainment of the perceiver, the perception of the other's stable characteristics will proceed from a different premise than if the behavior has no such relevant consequences. We may refer to such premises as *inferential sets* and assume that in the former case, when another's behavior has consequences for the perceiver, a *value-maintenance* set has been aroused. The consequent perceptions of the other's characteristics will then vary greatly as a function of whether the other's behavior promotes or interferes with goal attainment or the maintenance of values. If the other's behavior has no such relevant consequences even though he behaves in an objectively identical way, he will be perceived and evaluated in a different, more neutral, fashion.

Before stating the propositions which follow from these considerations, it is necessary to describe briefly the experimental procedure common to both experiments to be reported below. Groups of five to six subjects worked in parallel on a series of problem-solving tasks. A criterion of over-all success or failure was stated at the outset, and one member of the group (by prearrangement) failed to meet this criterion. For half the groups, the failure of this single member meant that he alone did not receive the promised reward. For the remaining groups, his failure meant that no one in the group received the reward. Trait ratings of this critical member were obtained before and after the tasks. The general proposition to be tested may now be stated as follows: since the value-maintenance set will be promoted to a greater extent when the failure of one affects the reward attainment of all, more negative characteristics will be ascribed to the failing person when his failure implies group failure than when it does not. More specific predictions will be stated after the procedure of Experiment I is described in greater detail.

¹This study was conducted in the program of the Organization Research Group of the Institute for Research in Social Science at the University of North Carolina. The O.R.G. is supported jointly by the Ford Foundation and the Office of Naval Research [Nonr-855(04)]. This research was partially financed by a grant to the senior author from the Duke University Research Council.

EXPERIMENT I

Procedure

Groups of five or six subjects, with one exception volunteers from the Introductory Psychology Course, met for an experiment on "group problem solving." The one exception in each group was a role-playing confederate who was carefully instructed to behave in a neutral fashion and to fail a standard pattern of tasks. In an introductory period prior to the main part of the experiment, procedures (including a "quiz-program" type task) and instructions were introduced which attempted: (a) to arouse the motivation of the subjects to do their best on the tasks required, (b) to encourage enough interaction so that subjects would have some basis for their initial ratings, and (c) to give the confederate a chance to establish himself initially as quite intelligent.

The subjects were told that the group problem-solving tasks were well-established measures of individual intelligence. Following the orientation procedures, the *before measure* (trait ratings) was introduced as a measure of intelligence in "that most neglected area, social sensitivity." Rating sheets were passed out containing 21 trait pairs such as friendly—unfriendly, lazy—energetic, etc., separated by a row of eight enclosed boxes, representing the scale points to be checked by each subject. Each subject rated the crucial member (the failing confederate) and one other naive subject without knowing that all others were also rating the confederate.

After the initial ratings had been completed, the experimenter told the subjects that the criterion of individual success and failure on the four tasks would be based on norms established at a rival college. The experimental groups were told that in order for the group to be successful and for any of the members to receive a monetary reward (\$1.00), *all* would have to attain the stated criterion of "passing" three of the four tasks. If one failed, all failed, and no one received the reward. The control groups were told that any individual in the group who attained the (same) criterion would receive the reward, regardless of how well or poorly the other subjects did.

The four tasks involved various verbal, arithmetical, and spatial skills all of which had a plausible relation to intelligence. In every group, the confederate was the only member to fail two of the four tasks. Special deception procedures were devised to allay the suspicions of the naive subjects, and the genuineness of the confederate's failure was uniformly accepted. There were generally observable differences between experimental and control groups in the extent to which the suspense mounted as the last task neared completion. Sometimes, in the "common fate" treatment, a subject or two would act as cheer leaders, urging the confederate through the items in hopes that their encouragement would bring success to the subject and hence to the group.

After the tasks were completed, the experimenter announced either that the group (experimental condition) or one member (control condition) would not receive the reward, and subjects were asked to rerate the same two members that they had previously rated.

The Motivation Variable. An attempt was made to vary the importance of "goal attainment" by preselecting subjects in terms of their achievement motivation following procedures described by McClelland *et al.* (5). Experimental and control groups were each composed of ten "highs" and ten "lows" (falling above and below the general class median respectively). Thus the experimental design was essentially a two-by-two design with ten subjects in each of four cells.

Hypotheses and Results

The main hypothesis of this first experiment is that the evaluation of the confederate will change more in a negative direction under experimental than control conditions, since his behavior prevents the subjects' goal attainment under the former and not the latter conditions.

A second hypothesis, which might be stated as a corollary of the first, is that the higher the need for achievement, the greater the difference between experimental and control conditions.

It is now important to consider the nature of the dependent variable, the rating scale. Twenty-one trait pairs were chosen to represent three *a priori* clusters: *perceived competence*, *perceived motivation*, and *likeability*. In addition to traits judged to fall into each of these clusters, there were several "beta traits" (cf.4) used as fillers to inhibit an evaluative response set. The critical trait, *dependable*, was included to be evaluated on its own. Table 1 shows the extent to which the average change on each cluster (plus *dependable*) differ reliably from zero.

It is apparent from a glance at the table that the average subject, regardless of experimental conditions, sees the confederate as less competent and less highly motivated after he has failed the tasks. While there are no differences between the conditions with regard to perceived motivation, there is a significantly greater decline in perceived competence for the experimental groups than for the control groups ($p < .05$ by analysis of variance). This specific difference was unexpected and was not replicated in the second experiment to be described below. It is difficult to interpret this finding unless the competence cluster data are viewed as an expression of evaluative change through the set of ratings judged by the subject to be most appropriate to the situation.

The most direct test of the main hypothesis involves the *likeability* cluster. The hypothesis is not supported by the data, since there is not a significant difference between experimental and control groups. Moreover,

TABLE 1
Experiment I: Mean Change Scores for Each Trait Cluster by Conditions

	Group			
	Experimental		Control	
	High n Ach	Low n Ach	High n Ach	Low n Ach
Competence Cluster (Competent, intelligent, quick-witted and well-organized)	1.60*	1.65*	.88*	.60*
Motivation Cluster (Competitive, hard-driving, conscientious, energetic, does his best)	1.02*	.96*	.98*	.38*
Likeability (Friendly, very likeable, modest, sympathetic toward others, pleasant, considerate)	.10	.52*	.12	.28
Dependable	2.10*	1.80*	1.20*	1.00*

Note.—The higher the score, the greater the mean change toward negative evaluation.

* $p < .05$ (significantly greater than zero).

the data actually contradict the hypothesis with regard to the variable of achievement motivation. Subjects who are low in achievement motivation show a greater negative change on likeability items than those high in achievement motivation. It appears that subjects with high achievement motivation inhibit negative evaluations of the confederate, possibly because his failure serves to emphasize their own success by contrast.

Since previous research has shown that "dependability" is a trait which is especially sensitive to changes in interpersonal perception (cf.2), it was hypothesized that the confederate will be seen as declining in dependability more under the experimental than the control situation. It was further predicted that this change would be greater for those subjects high in achievement motivation. The first part of this hypothesis was clearly supported ($p < .02$ by analysis of variance), but again the achievement motivation variable failed to discriminate.

Thus while the confederate's failure under conditions of common group fate does not result in his being rated down on the likeability cluster, he is seen to be significantly less dependable than the same person behaving in the same way under conditions of individual reward attainment.

As a final over-all index of evaluative change, an *a priori halo cluster* was constructed by combining the 14 most clearly evaluative traits. By placing each subject's "halo score" in a two-by-two factorial design, it was possible to test the main hypothesis (of greater change under experimental conditions) in terms of global evaluation. The hypothesis was confirmed at the

.05 level. Again the difference between highs and lows on need achievement was negligible.

Discussion of Experiment I

If one compares the reactions to a role-playing confederate under the experimental and control conditions of the first experiment, the results indicate that when the confederate's behavior is relevant to (i.e., prevents) the goal attainment of the perceiver, he is seen as relatively less competent, less dependable, and the over-all evaluation of him is more negative in tone. However, a group of traits specifically chosen to reflect likeability did not discriminate between subjects in the different experimental conditions.

There are several possible reasons for the failure of this particular prediction. It may be that a general reluctance to commit oneself to a "negative likeability" judgment is a powerful obscurant in the present situation, and that the culturally grounded inhibitions behind this reluctance operate less in other areas of interpersonal evaluation. This would help explain the significant results with regard to the competence and halo clusters which may serve, in a sense, as surrogates for a likeability judgment.

Of somewhat greater theoretical interest is the possibility that the primary hypothesis needs to be qualified to take into account certain perceptual preconditions. It is quite conceivable that the perceived *reason* for the confederate's failure plays an important role in any concurrent evaluation of him. Thus, a lowered estimate of the confederate's competence would lead to a different evaluative outcome than a similar reappraisal of the confederate's motivation. While we had hoped to promote the latter judgment by having the confederate initially appear quite intelligent, the results clearly show a decline on both attributes, for the change scores of the competence and motivation clusters are correlated to a significant extent (r for the experimental groups being $+.73$, for the control groups being $+.47$).

But, if the confederate is perceived to fail as a function of inadequate ability, there should be less of a tendency to evaluate him negatively than if his failure is motivational in origin. One's competence, by and large, is an attribute cluster over which the individual has little control—the sources of competence are in a sense external to the sphere of individual autonomy and responsibility. Motivation, on the other hand, is an attribute cluster likely to be perceived as self-caused or internally controlled by the behavior in question. Such considerations stem quite directly from the recent experiments of Thibaut and Riecken (7), which in turn exploded various implications of the work of Heider (1) and Michotte (6).

Examination of the correlation between perceived change in motivation and perceived change in likeability provides some support for this inter-

pretation. While this correlation is negligible ($r = +.02$) for the control subjects, as we would expect, the correlation rises to $+.36$ for the experimental subjects ($p = .10$). Thus it appears that there *is* some change in direct "likeability" evaluation under conditions of common group fate, when the confederate is seen as less motivated in the after than in the before ratings.

Since this relationship is a tenuous one, a second experiment was designed in an attempt to obtain greater control over the subject's attribution of motivational versus competence change. It was also considered highly desirable to provide a direct replication of the positive findings of the first experiment.

EXPERIMENT II

In planning the second experiment, an attempt was made to preserve the general procedure outlined above unless there were sound reasons for modification. The major change involved differential instructions designed to induce a perceived decline either in the confederate's motivation or his competence. For half the groups the instructions repeatedly stressed the fact that the tasks were designed to measure motivation ("how hard you're willing to work for an incentive") and that they had been shown in previous work to be unrelated to intelligence or aptitude. In the remaining groups, the high relationship with standard intelligence tests was stressed. A number of minor variations were also introduced. These are listed below.

a. The experiment was conducted by a different experimenter, with a different confederate, and with subjects from a different school.

b. A coding task was substituted for one which made more clear-cut intellectual demands on the subjects.

c. Achievement motivation was not a measured variable in Experiment II.

d. Each subject was asked to rate all other subjects on the before and after ratings. In order to reduce the tedium of repeated ratings, 10 of the original 21 traits were selected to represent the same trait clusters as in Experiment I. "Dependability" was also included.

In summary, this experiment was much like the first except that half the groups were told that the tasks definitely measured intelligence, the other half that they definitely did not measure intelligence. The design is again a two-by-two (experimental-control reward structure, and motivation-competence instructions) with ten subjects in each cell.

Results

Effectiveness of Experimental Manipulation. The first relevant question concerns the extent to which the instructions were successful in inducing

TABLE 2

Experiment II: Mean Change Score for Each Trait Cluster by Conditions

	Group			
	Experimental		Control	
	Motivation Instructions	Competence Instructions	Motivation Instructions	Competence Instructions
Competence Cluster (Competent, intelligent, quick-witted)	2.05*	1.90*	1.13*	2.00*
Motivation Cluster (Conscientious, does his best)	2.00*	1.30*	1.15*	1.95*
Likeability Cluster (Very likeable, considerate of others, pleasant, modest)	.23	.00	-.15	-.13
Dependable	3.00*	1.20*	.20	1.10*

Note.—The higher the score, the greater the mean change toward negative evaluation.

* $p < .05$ (significantly greater than zero).

the set to perceive failure as attributable to low motivation *or* lack of competence. When asked at the close of the experiment to rate the effectiveness of the tasks in tapping (a) motivation and (b) competence, the subjects generally responded in terms of the treatment to which their group had been assigned. Furthermore, subjects in the competence condition actually *perform* significantly better on the tasks than subjects in the motivation condition. (By analysis of covariance it was shown that a subject's ratings were not affected by his own performance.) However, as Table 2 shows, the motivation versus competence instructions did not affect the subjects' ratings on the relevant motivation and competence clusters. In other words, while the subjects perceived the tasks as measures of motivation or competence in line with the instructions, this had little effect on whether the confederate was seen as failing for reasons of motivation or ability.

Changes in Evaluation of the Confederate. While we were not entirely successful in creating the conditions for a crucial test of the revised hypothesis, it is still meaningful to examine the changes in evaluation which did take place. Turning to the *likeability cluster*, Table 2 indicates that the results are in the right direction, but the predicted interaction between the experimental-control conditions and the motivation-competence conditions falls far short of significance. This remains true when the discrepancy between a subject's rating of the confederate and his rating of the other subjects is used as the adjusted measure.

TABLE 3
Means and Variances, Halo Cluster, Experiment II

	Groups			
	Experimental		Control	
	Motivation Instructions	Competence Instructions	Motivation Instructions	Competence Instructions
\bar{X}^*	23.0	17.3	14.4	18.9
s^2	114.00	39.34	4.49	90.77

* The greater the mean change, the more negative the "after" evaluation.

An a priori *halo cluster* was again constructed to give a test of general positive evaluative change. This type of measure yielded a significant difference between experimental and control conditions in Experiment I. It was here predicted that there would be more negative halo in the motivation-experimental condition than in any other combined condition. While this prediction is borne out by inspection, and the interaction is large enough to yield a significant F ratio under normal circumstances, marked heterogeneity of variance between subconditions makes the conventional F-test inapplicable. The means and variances are presented in Table 3. It is clear that the motivation-experimental subgroup has the largest mean, but there is wide variability among the individual scores composing this mean. While some subjects in the motivation-experimental subgroup changed markedly in a negative direction, a few changed very slightly. In the motivation-control subgroup, however, the change was uniform and moderate.

In Experiment I, subjects in the experimental groups saw the confederate as declining more in *dependability* than did subjects in the control groups. In line with the revised hypothesis of Experiment II, this decline should be greater when the reason for failure is perceived to be motivational than when it is perceived to be a matter of competence. Table 2 shows that this is indeed the case. The F ratio reflecting the interaction between experimental-control and motivation-competence conditions is highly significant ($F = 7.45$, $d.f. = 1$ and 36). Thus, despite the fact that the two sets of instructions did not create differential perceptions of the confederate's motivation, they did have the predicted effect on the outcome ratings of dependability.

The main (revised) hypothesis of the present experiment assumes that there will be a greater decline in positive evaluation under experimental than under control conditions if the confederate's failure is attributed to lack of motivation. Therefore, we would expect the correlation between

perceived change in motivation and perceived change in evaluation to be more positive under experimental than under control conditions. That is, the person with low motivation will be liked less under experimental than under control conditions. The correlations using the data from Experiment II were $+.03$ for the experimental subjects and $-.21$ for the control subjects. For Experiment I on the *same* traits (the Motivation and Likeability clusters) the correlations were $+.03$ for the experimental subjects and $-.28$ for the control subjects. None of these correlations is significant, nor does either experimental correlation differ significantly from either control correlation. It is encouraging, however, that the differences are in the predicted direction in both experiments and that they are very similar in magnitude.

It is somewhat surprising that the experimental group correlation is approximately zero and the control group correlation is negative. Apparently, when "conscientious" and "does his best" are employed as an index of motivation, the control subjects like the confederate when he is perceived to be poorly motivated *better* than when he is perceived to be highly motivated. Since the two motivation traits seem to be positively evaluative, this correlation is even more interesting and suggests the presence of competitive feeling in the control groups.

GENERAL DISCUSSION

There is evidence from both experiments that a person who behaves in a certain way will be perceived differently as a function of the relevance of his behavior to the value maintenance of the perceiver. Specifically, in both experiments, the failing confederate is perceived to be less dependable when his failure prevents others from reward attainment. Furthermore, when the tasks are presented as measures of motivation in Experiment II, this decline in perceived dependability is clearly greater than when the tasks are presented as an intelligence test. By inference, the degree to which the confederate is seen as responsible for behavior which causes the group to fail is a definite factor in evaluating his dependability. The correlational data bear this out in a more tentative way with regard to his "likeability." Objectively, this contingency between change in perceived motivation and change in evaluation is not entirely reasonable because an incompetent but highly motivated person can be equally obstructive as far as goal attainment is concerned. Subjectively, however, when the locus of phenomenal causality is perceived as internal to the agent of frustration, negative evaluation is more severe.

These related experiments may be seen as a demonstration of the effects of set on the perception of the characteristics of others. Different inferences

will be drawn from the same behavior, as a function of the set promoted by the structure of the social situation. The behavior of others does not appear to have a constant meaning, and the attribution of stable characteristics to the behavior is dependent on the significance of his behavior for the perceiver's own value-maintenance of goal attainment.

SUMMARY

Two experiments were conducted in an effort to demonstrate that different inferences will be drawn from the same behavior when this behavior does or does not have personal relevance for the perceiver. A common experimental procedure was adopted for both experiments which involved the ultimate failure of a confederate to meet the announced norms on a series of tasks. The personal relevance of this failure was varied by initially announcing that the attainment of a monetary reward would either be on a group or "common fate" basis (experimental groups) or on an individual achievement basis (control groups). In the first experiment, subjects were preselected in terms of their general achievement motivation, in an attempt to vary the importance of attaining the reward. In the second experiment, achievement motivation was not a measured variable, but the tasks were presented as either measures of motivation or intelligence. The main hypotheses and results of both experiments may be summarized as follows:

1. When the reward attainment of each subject depends on the successful performance of all subjects, rather than being determined on an individual basis, (a) a failing subject will be seen as relatively less likeable, (b) across all evaluative traits there will be a greater negative halo, and (c) the failing subject will be seen as less dependable in the former than in the latter condition. The results show that (a) is not supported by either experiment, (b) is supported by Experiment I, and as predicted in Experiment II under "motivation instructions" (although heterogeneity of variance precludes a rigorous test), and (c) is supported by both experiments. It is important to note that both negative halo increase and dependability decline are greater when the tasks are presented as measures of motivation in the second experiment —i.e., when the locus of causality is internal to the subject who fails.

2. There will be a higher correlation between perceived motivation decline and likeability decline under experimental than control conditions. The results show that, while the correlations in neither experiment differ significantly from each other, they are both in the predicted direction.

While the results of the present investigations are generally in line with theoretical expectations, it is clear that there are many difficulties in controlling social stimulus cues to yield uniform and consistent patterns of inference. Nevertheless, the results of the present research seem to encour-

age the belief that progress in understanding the processes of social perception can be served by such experimental attempts.

Manuscript received: July 30, 1966

Revised manuscript received: October 10, 1966

Edward E. Jones

Organization Research Group

University of North Carolina

Chapel Hill, North Carolina

REFERENCES

1. Heider, F., "Social Perception and Phenomenal Causality," *Psychological Review*, 1944, 51, 358-374.
2. Jones, E. E., "Authoritarianism as a Determinant of First-Impression Formation," *Journal of Personality*, 1954, 23, 107-127.
3. Jones, E. E., "Inferential Sets in Social Perception," paper read at the meeting of the American Psychological Association, San Francisco, 1955.
4. Lemann, T. B., and R. L. Solomon, "Group Characteristics as Revealed in Sociometric Patterns and Personality Ratings," *Sociometry*, 1952, 15, 7-90.
5. McClelland, D. C., J. W. Atkinson, and E. L. Lowell, *The Achievement Motive*, New York: Appleton-Century-Crofts, 1953.
6. Michotte, A., "La Perception de la Causalité," Louvain, 1946.
7. Thibaut, J. W., and H. W. Riecken, "Some Determinants and Consequences of the Perception of Social Causality," *Journal of Personality*, 1955, 24, 113-133.



